DOCUMENT RESUME

ED 370 413 FL 022 191

| | |
|---|---|
| AUTHOR | Roby, Linda M., Ed. |
| TITLE | Kansas Working Papers in Linguistics. Volume 19. |
| INSTITUTION | Kansas Univ., Lawrence. Dept. of Linguistics. |
| PUB DATE | 94 |
| NOTE | 475p.; For individual papers, see FL 022 192-207. |
| AVAILABLE FROM | Editors, KWPL, LGSA, Linguistics Department, 427 Blake Hall, University of Kansas, Lawrence, KS 66045. |
| PUB TYPE | Collected Works - Serials (022) |
| JOURNAL CIT | Kansas Working Papers in Linguistics; v19 n1-2 1994 |
| EDRS PRICE | MF01/PC19 Plus Postage. |
| DESCRIPTORS | African Languages; American Indian Languages; Chinese; Dakota; *Descriptive Linguistics; English (Second Language); Hausa; Higher Education; Korean; *Language Research; Language Typology; *Linguistics; Native Speakers; Semantics; Sentence Structure; *Structural Analysis (Linguistics); Uncommonly Taught Languages |
| IDENTIFIERS | Clitics; Kansa; Karankawa; Xhosa |

ABSTRACT

This collection of papers presents the latest original research by the institutions. The papers in Number 1 are: (1) "Xhosa departments of the University of Kansas, as well as contributors from other institutions. The papers in Number 1 are: (1) "Xhosa Nominal Tonology: A Domain-Based Approach" (Mbulelo Jokweni); (2) "On the Representation of Mandarin Syllable Structure" (Feng-Lan Kuo); (3) "Vendler Classes and Reinterpretation" (Michihiko Kawamura); (4) "Negative Polarity Items and the Semantics of the Particles '-to' and '-na' in Korean" (Yae-sheik Lee); (5) "Clitics, Case Checking, and Causative Constructions" (Xavier Villalba); (6) "The Dual Status of the Null Object in Chinese" (Yanfeng Qu); (7) "On the Orientation Problem in Korean 'CAKI' Binding and the Typology of X Reflexive Binding" (Mi-Hui Cho); (8) "Complementation of Hausa Aspectual Verbs" (Lawan Danladi Yalwa); (9) "Deriving the Distribution of Conjunctions" (Ed Zoerner); and (10) "L2 Acquisition of English Reflexives by Native Speakers of Korean" (Hye-Ryun Kim). All six papers in Number 2 deal with Native American Languages: (11) "Karankawa Linguistic Materials" (Anthony P. Grant); (12) "Word-Level Nominalization in Choctaw" (Marcia Haag); (13) Reflexives in Mohawk" (Nancy Bonvillain); (14) "Proto-Algic V: Doublets and Their Implications" (Paul Proulx); (15) "The Limit of Structure Preservation in Dakhota Lexical Phonology" (John Kyle); and (16) "Notes on the Kansa Word List of Maximilian, Prince of Wied" (Robert L. Rankin). Each paper is followed by a reference list. (MDM)

***********************************************************************
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
***********************************************************************

# Kansas Working Papers in Linguistics

Volume 19, Number 1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. Manuel Dupont
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 370 413

FL022191

*Kansas Working Papers in Linguistics* (*KWPL*) is a regular publication of the Linguistics Graduate Student Association, Department of Linguistics, University of Kansas, Lawrence, Kansas, 66045.

Aim: *Kansas Working Papers in Linguistics* is intended as a forum for the presentation, in print, of the latest original research by the faculty and students of the Department of Linguistics and other related departments at the University of Kansas. Papers contributed by persons not associated with the University of Kansas are also welcomed. The papers published in *KWPL* may not be reproduced without written permission from the Linguistics Graduate Student Association.

Calls for papers usually end in February of each year. Check yearly Call for Papers for exact date. Please send manuscripts, requests for individual orders, and inquiries to:

Editors, *KWPL*
LGSA, Linguistics Department
427 Blake Hall
University of Kansas
Lawrence, Kansas 66045
E-mail: LGSA@kuhub.cc.ukans.edu

The price list for previously published volumes is available upon request. Postage is paid for requests from the United States and Canada. For orders outside the United States and Canada, please add US$3.00 per volume to help defray the costs of postage. A cumulative index may be found in Volume 16 and is available upon request. All institutions with a similar publication are encouraged to request a reciprocal agreement. Standing orders for Number 1, Number 2 (SNAL) or both are welcomed.

The editor would like to express her sincere appreciation to the faculty and staff of the Department of Linguistics for their assistance and encouragement, especially Professor Sara T. Rosen, Mary S. Linn, and Giulia R.M. Oliverio. I would also like to thank the Graduate Student Council for their continuing support.

Cover design, with a minor botanical alteration, by David Andrew Toshraah Nokose Skeeter.

# Kansas Working Papers in Linguistics

edited by

Linda M. Roby

Volume 19, Number 1
1994

Partial funding for this journal is provided by the Graduate Student Council
through the Student Activity Fee.

© Linguistics Graduate Student Association
University of Kansas, 1994

**Kansas Working Papers in Linguistics**
**Volume 19, Number 1**
**1994**

Xhosa Nominal Tonology: A Domain-Based Approach
Mbulelo Jokweni....1

On the Representation of Mandarin Syllable Structure
Feng-Lan Kuo....19

Vendler Classes and Reinterpretation
Michihiko Kawamura....53

Negative Polarity Items and the Semantics of the Particles *-to* and *-na* in Korean
Yae-sheik Lee....89

Clitics, Case Checking, and Causative Constructions
Xavier Villalba....125

The Dual Status of the Null Object in Chinese
Yanfeng Qu....149

On the Orientation Problem in Korean *caki* Binding and the Typology of X° Reflexive Binding
Mi-Hui Cho....165

Complementation of Hausa Aspectual Verbs
Lawàn Danlādì Yàlwa....185

Deriving the Distribution of Conjunctions
Ed Zoerner....217

L2 Acquisition of English Reflexives by Native Speakers of Korean
Hye-Ryun Kim....235

# XHOSA NOMINAL TONOLOGY: A Domain-based Approach

Mbulelo Jokweni
University of Illinois

Abstract: In this paper I argue that a domain-based approach explains in a straight forward way, by means of a single H tone spread rule, the complex tonal structure of isiXhosa nouns. For every noun type I propose H tone-motivated domain structures, and I call them Tone Domains (TDs). The number of TDs is determined by the number of lexical H tones in a word. I then demonstrate that the H tone in each TD spreads to its right edge provided the Obligatory Contour Principle (OCP) is not violated. The significance of this approach is that it explains why the preprefixal H tone would spread to the penult in some words, but to the antepenult in others. Also, it explains why the preprefixal H tone is subjected to the rule in some nouns, but not others.

## 1. Introduction.

In this paper I propose and discuss a domain-based approach to Xhosa nominal tonology. I show that the domain-based approach to phonology, developed by Kisseberth (1992), is a successful theoretical device for the analysis of complex tonal systems such as Xhosa, a Bantu language spoken in South Africa. In particular, I argue that within this framework there is only one H tone spread rule in Xhosa nominals, and that this rule applies within a tone domain which is triggered by the presence of a H tone.

Kisseberth (1992) defines a domain as 'a sequence of phonological material enclosed by a left and right bracket'. He asserts that this approach assumes the existence of phonological rules, and that 'prior to the application of a rule to a representation, a domain structure should be assigned to the representation.'

I must from the onset state that although this approach entails the same central idea of 'prosodic domains' developed by Selkirk (1980a, b) and Nespor and Vogel (1986), it nonetheless represents a somewhat different view in that it does not rely on the hierarchy of prosodic levels. For example, my analysis will show that a word may have a sequence of unlayered tonal domains depending on the number of lexical H tones.

In general Kisseberth takes the position that the domain types are reflexes of phonological, morphological, and syntactic structure. I should also mention that in addition to H tones, the morphological structure also plays a role in assigning domains for the phonological representations. Evidence for the projection of domains on the basis of morphology is found in 'Cole 1992'.

The outline of this paper is as follows. First, I briefly describe in section 2 some morphological, phonological and tonal phenomena which are relevant to the assignment of domains and the application of phonological processes. In section 3 I present some data from Xhosa and briefly outline the problems of tonal

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 1-17

analysis. In section 4 I discuss the assignment of domain structures and the application of phonological processes within domains. Finally, in section 5 I summarize the main points of my analysis.

### 2. Morphological, phonological and tonal structure.

The morphological structure of Xhosa nominals is not simple. It is particularly made complex by the 'irregular' behavior of the noun class prefix. As shown in 1, a Xhosa noun is composed of the noun class prefix and the stem. The prefix has an initial vowel which is referred to as the preprefix in the literature (Pahl 1976). The preprefix is a copy of the prefix vowel.

| | | | | | | |
|---|---|---|---|---|---|---|
| (1) | Class 3 | *ilí-zwe* | 'country' | Class 6 | *amá-zwe* | 'countries' |
| | Class 7 | *isí-zwe* | 'nation' | Class 8 | *izí-zwe* | 'nations' |

In classes 1 and 3 the vowel of the proto-Bantu prefix **mu-* is lost and the nasal consonant is syllabic (2).

| | | | | | | |
|---|---|---|---|---|---|---|
| (2) | Class 1 | *um´-ntu* | 'person' | Class 2 | *abá-ntu* | 'people' |
| | Class 3 | *úm-thí* | 'tree' | Class 4 | *ími-thí* | 'trees' |

Similarly, the vowel of the proto-Bantu class 9 prefix **ni-* is lost, but unlike classes 1 and 3 the remaining nasal is not syllabified. Instead it is pronounced homorganically with the consonant of the following stem syllable (3).

| | | | |
|---|---|---|---|
| (3) | Class 9 | *ín-taba* | 'mountain' |
| | | *ín-komó* | 'cow' |

In classes 5 and 11 the whole prefix deletes if the stem has two or more syllables (4).

| | | | | | | |
|---|---|---|---|---|---|---|
| (4) | Class 5 | *ilí-zwe* | 'country' | but | *í-háshe* | 'horse' |
| | Class 11 | *úlu-thí* | 'stick' | but | *ú-dongá* | 'wall' |

Finally, in class 5 the prefix is optionally dropped out if the stem is monosyllabic (5).

| | | | |
|---|---|---|---|
| (5) | Class 5 | *ilí-zwe* or *í-zwe* | 'country' |
| | | *íli-zwí* or *î-zwí* | 'voice |

In general nominal stems do not have more than three syllables. Nevertheless longer stems are attained by extensions such as the dimunitive suffix *-ana*, augmentative suffix *-kazi*, feminine suffix *-kazi*, and locative suffix *-eni* (6).

| | | | | | |
|---|---|---|---|---|---|
| (6) | *amá-Xhosa* | 'Xhosa' | : | *ama-Xhosákazi* | 'Xhosa women' |
| | *amá-Xhosa* | 'Xhosaa' | : | *ema-Xhóseni* | 'to the Xhosas' |
| | *ín-taba* | 'mountain' | : | *in-tátyana* | 'small mountain' |
| | *ín-taba* | 'mountain' | : | *in-tabákazi* | 'huge mountain' |

The affixation of the suffixes involves some morpho-phonological processes which are not relevant to the issues discussed in this paper, hence I am not going to discuss them here. Suffice to mention that these suffixes, together with the prefix in the case of the locative, do not contribute any tone to the noun.

Nouns (or all word categories) in phrase final position have a long penultimate vowel. The same phenomenon is exhibited by the other closely related languages such as Zulu and Swati. However, there is no vowel length contrast between words in these languages. It is thus assumed that the long penultimate vowel is not underlying, but is derived by a vowel lengthening rule (Clark 1988).

Xhosa nouns are divided into two tonal groups: L nouns and H nouns. This division is based on the claim that L nouns do not have an underlying H tone on any of their stem syllables, and that the H tone which surfaces in long forms is contributed by the preprefix. The diagnostic for L nouns and H nouns is the absence and presence of the H tone in the short nominal form. A short form is without the preprefix and it would appear in phrases like: 'Akúkho ... ' i.e. ' 'There is no ... ' (7).

| | | | | | |
|---|---|---|---|---|---|
| (7) | a. | L nouns: | *Akúkho* | *zwe* | 'country' |
| | | | | *ba-ntu* | 'people' |
| | | | | *ba-ntwana* | 'children' |
| | | | | *si-bonjana* | 'small pole' |
| | | | | *ma-Xhosakazi* | 'Xhosa women' |
| | | | | *ma-dangatyekazi* | 'big flames' |
| | b. | H nouns: | *Akúkho* | *ma-zwí* | 'voices' |
| | | | | *ma-gwalá* | 'cowards' |
| | | | | *ma-háshe* | 'horses' |
| | | | | *ba-fâzi* | 'women' |
| | | | | *m-hédeni* | 'heaten' |
| | | | | *ma-Gcaléka* | 'people of Gcaleka clan' |
| | | | | *m-nyhádalá* | 'game' |
| | | | | *m-babalá* | 'old buck' |
| | | | | *m-phongôlo* | 'cask' |
| | | | | *bhóbh-óbyi* | 'African hoopoe' |
| | | | | *m-dlámbîla* | 'rock cobra' |

Notice that in (7a) both the prefix and the stem are toneless. When the preprefix is used all the forms given in (7a) surface with a H tone, e.g. *í-zwe, abá-ntu, abá-ntwana, isi-bónjana, ama-Xhosákazi, ama-dangatyékazi.* Thus, in the absence of evidence for the underlying existence of a H tone in these nouns it is construed that the H tone originates from the preprefix. The claim that the preprefix has an underlying H tone finds further support from monosyllabic and bisyllabic L nouns without the prefix. As shown in (8) the preprefix in these nouns is invariably H toned.

| | | |
|---|---|---|
| (8) | *í-zwe* | 'country' |
| | *í-gusha* | 'sheep' |
| | *ín-taba* | 'mountain' |

The nouns given in (7b), on the other hand, represent a complex underlying tonal structure of the stem, namely H, OH, HO, HH, HOO, OHO, HOH, OOH, OHH, and HHH. (letters H and O represent the presence and the absence of a H tone, respectively.) Although the patterns represented as underlying HH# surface with a falling tone on the penult, I will show in the subsequent discussion that some phonological processes apply to yield the falling tone on the penult and the deletion of the H tone on the final syllable. I will also show that in these forms the H tone is multiply linked to the tone bearing units which are moras. In the long form the nouns in (7b) surface with a H tone on the prefix or preprefix (9).

| | | |
|---|---|---|
| (9) | *áma-zwí* | 'voices |
| | *amá-gwalá* | 'cowards' |
| | *áma-háshe* | 'horses', etc. |

From what we have seen in the long forms I claim that there is a constrained spread of the lexical H tone. The H tone spreads to the right and delinks from the left branches. The question now is: How can this spread rule be formalized? In *abá-ntu* 'people' the H tone spreads to the penult and in *am-Xhosákazi* it spreads to the antepenult. In *ama-háshe* 'horses' it does not spread while in *aná-gwalá* 'cowards' it spreads to the antepenult. I do not want to say that this spreading is arbitrary. Thus proposing a domain-based approach is an attempt to give a systematic account of the rule of H spread in Xhosa nominals. This approach will show that there is one H spread rule, and that the H tone spreads to the right edge of a domain (i.e. the rightmost mora in a domain).

Having described the pertinent morphological, phonological and tonal phenomena, I now turn to the tonal analysis and the problems associated with it.

## 3. Problems of tonal analysis.

In order to maintain the principle of H spread it is important to identify the target tone bearing units since the H tone does not spread all the way to the final syllable. One alternative is to posit two H spread rules. The first rule spreads the H tone to the penult if the stem is monosyllabic (10).

```
(10)  H
      | \
      aba-ntu
```

The second rule spreads the H tone to the antepenult if the stem is longer (11).

```
(11)  H
      | \   \ \
      ama-Xhosakazi
```

This proposal is falsified by examples such as *ím-babalá* 'old buck'. In *im-babalá* the H tone does not spread to the antepenult whereas it does in *amá-gwalá* 'cowards. Thus the facts shown by *ím-babalá* and *amá-gwalá* can lead to a conclusion that a H tone simply spreads to the prefix if the noun is H and to the antepenult if the noun is L. This proposal would also fail to account for the H spread to the penult in monosyllabic L nouns such as *abá-ntu* 'people'. Above all we notice that the tonal analysis of these data involves random counting of syllables. Note that there is no motivation for the H spread to the penult in some nouns and the spread to the antepenult in others.

We notice that in both *ím-babalá* and *amá-gwalá* the H tone spread does not cross the prefix boundary. Taken at face value, the failure of spreading in *ím-babalá* could be easily ascribed to the 'irregular' nature of the prefix. However, the locative plural form has a 'regular' prefix yet the spread does not cross the prefix boundary (12).

(12) H H
| \ |
ezim-babaleni

This state of affairs counts as motivation for including the role of morphology in the assignment of domain structures.

## 4. Domain-structure assignment and application of rules.

Evidence from all nouns show that the final syllable is not a target for H spread. A second fact shown by L nouns with multi-syllabic stems is that the penult is not generally affected by the H tone spread rule. For convenience the examples which show these facts are repeated in (13).

| | | |
|---|---|---|
| (13) | *abá-ntu* | 'people' |
| | *abá-ntwana* | 'children' |
| | *isi-bónjana* | 'small people' |
| | *ama-Xhosákazi* | 'Xhosa women' |

On the basis of these facts I postulate the construction of two domain structures, namely word domain (WD) and tone domain (TD). First, I assign a WD structure which excludes the final toneless syllable provided it is not the only syllable of the stem. This exception is intended to avoid constructing a word domain structure which does not have a stem syllable. Then I assign a TD structure which is triggered by the presence of a H tone.

To illustrate WD structure assignment I use the underlying forms of the L nouns given in (13). First, the rightmost syllable of the word projects a Right bracket to its Left if the stem is multi-syllabic, otherwise the rightmost syllable projects a Right bracket to its Right. Then the Bracket Matching Convention (BMC) places a matching Left bracket to the Left of the initial syllable of the word (14a).

(14a) *abá-ntu* (mono-syllabic)
H
|
[aba-ntu]

*abá-ntwana* (multi-syllabic)
H
|
[aba-ntwa] na

*isi-bónjana* (multi-syllabic)
H
|
[isi-bonja] na

*ama-Xhosákazi* (multi-syllabic)
H
|
[ama-Xhosaka] zi

The phonological representation contained in a WD is not relevant for the application of the tonal rule of H tone spread. A WD simply lays the foundation for the construction of a TD which also exludes a final toneless syllable. The exclusion of a toneless final syllable form the TD is equivalent to the Final Projection in Xitsonga (Kisseberth 1992).

A TD is constructed inside a WD in the following way. The rightmost syllable of a WD projects a Right bracket to its Left and the BMC provides a matching Left bracket to the Left of the initial syllable of a WD. Thus in (14b) the outer brackets mark WDs and the inner brackets mark TDs.

(14) H
|
[ [aba-] ntu ]

H
|
[ [aba-] ntwa ] na

H
|
[ [isi-bo] nja] na

H
|
[ [ama-Xhosa] ka ] zi

The principles of domain-structure formation illustrated above apply generally to all nouns. The only different is that H nouns have multiple H tones. This situation necessitates the division of a TD into two or more smaller unlayered TDs. Thus the assignment of domain structures in H nouns is achieved by three steps. The first step assigns a WD structure. The second step assigns a TD structure, and finally the third step divides a TD into smaller TDs.

A WD structure in H nouns is obtained in the following way. The rightmost syllable of a noun projects a Right bracket to its Left if it is toneless, otherwise the rightmost syllable projects a Right bracket to its Right (16a). Then the BMC places a matching Left bracket to the left of the initial syllable.

(16a) *áma-zwí* 'voices' (H final syllable)

```
 H            H
 |            |
[ama   -    zwi]

 áma-háshe   'horses'          (Toneless final syllable)
 H           H
 |           |
[ama   -   ha] she

 úm-nyhádalá   'game'          (H final syllable)
 H             H        H
 |             |        |
[um   -   nyha    da   la]

 úm-hédeni    'heathen'       (Toneless final syllable)
 H            H
 |            |
[um     -   hede] ni
```

Also, to obtain a TD from the WD structures given in (16a), the rightmost syllable of a WD projects a Right bracket to its Left if it is toneless. If it is H it projects a Right bracket to its Right. Then the BMC provides a matching Left bracket to the Left of the initial syllable (16b).

```
(16b)
    H            H
    |            |
  [ [ama    -   zwi] ]

    H            H
    |            |
  [ [ama    -   ha] ] she

    H            H   H
    |            |   |
  [ [um    -   nyhadala] ]

    H            H
    |            |
  [ [um     -   he] de] ni
```

Finally, a TD with more than one H tone is divided into smaller TDs. The number of these smaller TDs is proportional to the number of H tones associated with the syllables in the main TD. To obtain these structures each H toned syllable must project a Left bracket to its Left. Then the BMC provides a matching Right bracket (16c).

```
(16c)
      H            H
      |            |
  [ [ [ama   -]  [ zwi] ] ]
```

```
    H           H
    |           |
[ [ [ama  -]  [ ha] ] ] she

    H            H        H
    |            |        |
[ [ [um  -]  [ nyhada]  [ la] ] ]

    H            H
    |            |
[ [ [um  -]    [ he] ]  [de] ni
```

From the surface forms given in (16) we notice that the H tone does not spread to the right edge of a domain. I ascribe the failure of the H tone spread in these nouns to an OCP (a constraint that two H tones cannot be adjacent). Notice that the same OCP constraint is responsible for the failure of a H spread to the edge in *áma-gwalá* 'cowards' (17)

(17) *áma-gwalá* 'cowards'

```
    H            H
    | \          |
[ [ [ama-gwa]  [ la] ] ]
```

Another alternative would be the postulation of a Pre-H Projection in (17). A Pre-H Projection is postulated by Kisseberth to account for similar cases in Xitsonga. This rule causes toneless syllable in front of a H-toned syllable to project a Left bracket to its Left.

According to this proposal *amá-gwalá* would have a domain-structure given in (18).

(18) *amá-gwalá*

```
    H                 H
    | \               |
[ [ [ama  - ] [gwa]  [ la] ] ]
```

In (18) the H tone does spread to the edge of a domain.

We also notice that in *ím-babalá* 'old buck' and *ez ím-babaléni* 'to the old bucks' the H tone of the preprefix does not only fail to spread to the edge of a domain, but does not go beyond the prefix (19).

(19) *ím-babalá* 'old bucks'

```
    H          H
    |          |
[ [ [im-baba] [la ] ] ]
```

*ezím-babaléni* 'to the old bucks'

```
    H          H
    |\         |
```

[ [ [ezim-baba] [le] ] ] ni

From what we see in (19) there is no reason why the H tone cannot spread to the first *-ba-*. On the basis of these facts I conclude that neither the OCP nor the Pre-H Projection is relevant to (19). I argue that in these nouns the H spread is blocked by a domain boundary which is projected by the stem. Specifically, the Left edge of a H stem projects a Left bracket to its Left (20). I must also mention that the stem is not always a relevant domain for the application of the H spread rule. It only blocks the H tone of the preprefix from crossing the stem boundary if there is a H tone on the stem.

```
(20)   H           H
       |           |
    [ [ [im -] [baba] [la] ] ]

       H            H
       |\           |
  [ [ [ezim-] [baba] [le] ] ] ni
```

Given a domain structure the easiest way of accounting for the H tone spread is to say that the H tone spreads to the right edge of a domain. According to Hyman (1990) processes which are restricted to the initial or final position of a domain are called domain-limit rules. Following Archangeli and Pulleyblank's (1992) framework, Hyman proposes the following parameters for the domain-limit rules (21).

(21) a. domain : PW, CG, PP, IP, U
b. edge : left, right
c. function : insert, delete, spread, delink
d. trigger : H, L, etc.
e. target, conditions : specific tone(s) and or tone-bearing units (TBUs)

For our purpose in this paper the parameters are as follows (22).

(22) a. domain : TD
b. edge : right
c. function : spread
d. trigger : H
e. target conditions : toneless mora

Subsequent to the H spread there is a Left Branch Delinking rule which delinks the left branches after spreading. These rules interact with other phonological processes such as the lengthening of the penultimate vowel. Finally all toneless vowels are assigned a L tone by default. All these rules interact in the manner represented in (23) to yield the surface forms.

```
(23)   abá-ntwana  'children,     abá-ntu  'people'
       H                                   H
       |                                   |
       aba-ntwana          U.R.            aba-ntu
```

```
   H                                   H
   |                                   |
[ [aba-] ntwa] na     D-Structure   [ [aba-] ntu]

   H                                   H
   |                                   |
[ [aba-] ntwaa] na    Lengthening   [ [abaa-] ntu]

   H                                   H
   | \                                 | \\
[ [aba-] ntwaa] na    Spreading     [ [abaa-] ntu]

   H                                   H
   | \                                 | \\
[ [aba-] ntwaa] na    Delinking     [ [abaa-] ntu]
[ [àbá-] ntwàà] nà    Low Default   [ [àbáá-] ntù]
[abántwaana]          S.R.            [abáántu]
```

As observed in *abá-ntu* (23) the Left branch Delinking rule does not apply to long vowels. It is blocked by a constraint given in (24).

```
(24)            * H
                 /   \
                μ     μ
                 \   /
                   &
```

This constraint finds support from the fact that a rising tone does not exist in Xhosa. Thus delinking in long vowels would yield incorrect surface forms such as * [abàántu].

The goal of domains is to put barriers or boundaries for the application of the H spread rule. This perhaps constitutes a universal rule for the theory of domains, namely that H tone spreading does not cross a domain boundary. Accordingly, in the representation given in (23) the H spread rule applies only within the tone domain. Anything outside this domain is invisible to this rule. More evidence for no-crossing of domain boundary comes from H nouns where there is a sequence of tone domains. The domain structure in *ím-babalá*, repeated in (25) confirms this point.

```
(25)  H             H
      |             |
   [ [ [im-] [baba] [la] ] ]
```

In (25) the stem domain boundary prevents the application of the H tone spread, hence there is no difference between the underlying and the surface tonal patterns.

Domains do not interfere with OCP. The first OCP effect is the H tone polarity exhibited by H nouns such as *ízi-kólo* 'schools', *áma-háshe* 'horses', etc.

These nouns have a sequence of tone domains, and the surface structure is derived in the manner shown in (26).

| | | |
|---|---|---|
| (26) | *ízi-kólo* | 'schools' |
| | U.R. | H H<br>\| \|<br>izi-kolo |
| | D-Structure | H H<br>\| \|<br>[ [ [izi-] [ko] ] ] lo |
| | Lengthening | H H<br>\| \|<br>[ [ [izi-] [koo] ] ] lo |
| | Spreading | H H<br>\| \| \\<br>[ [ [izi-] [koo ] ] ] lo |
| | Low Default | [ [ [ízì-] [kóó ] ] ] lò |
| | S.R. | [ízi-kóólo] |

In accordance with the formulation of the H spread rule we would expect the H tone linked to the preprefix in (26) to spread to the edge of the tone domain. The fact that this does not happen shows that the H spread rule respects the OCP. However, in the second tone domain the H tone does spread to the second mora. This is made possible by the fact that the final syllable is toneless.

Also, the same fact is shown by monosyllabic H nouns such as *î-zwí* 'voice' (27).

| | | |
|---|---|---|
| (27) | | |
| | U.R. | H H<br>\| \|<br>i - zwi |
| | D-Structure | H H<br>\| \|<br>[ [ [i -] [ zwi] ] ] |
| | Lengthening | H H<br>\| \|<br>[ [ [i i-] [ zwi] ] ] |
| | Spreading | N/A |
| | Low Default | [ [ [í ì -] [ zwí] ] ] |
| | S.R. | [í ì-] [ zwí] |

Upon lengthening the penultimate vowel, the preprefix H in (27) is expected to spread to the second mora to yield * [*íí-zwí* ] which is an incorrect surface form. Thus the correct surface form [*íî-zwí* ] derives from the failure of the H tone to spread to the second mora despite the fact that the following adjacent H tone is in a separate domain.

In Xhosa the depressor consonants (mainly voiced consonants) play an important role in shaping the tonal structure of words. For example, the plural form of the L noun *si-londa*, *isí-londa* 'wound' (short and long form) is *zi-londa*, *izi-lônda* 'wounds'. The surface form *isí-londa* is derived by the application of the H spread rule in the manner discussed above. However, the plural form of the same noun surfaces with a falling tone on the penult.

Tonologists generally attribute this situation to the effect of the depressor consonant on the H tone (Clark 1988), (Khumalo 1989), etc. They argue that the depressor consonant shifts the H tone to the following syllable provided the following syllable does not have a depressor consonant. This process is called the depressor shift. Thus, according to this proposal *izi-lônda* is derived in the following manner (28).

(28) *izi-londa* 'wounds'

```
                      H
                      |
U.R.                  izi-londa

                      H
                      |
D-Structure         [ [izi - ] lo] nda

                      H
                      |
Lengthening         [ [izi - ] loo] nda

                      H
                      |\
Spreading           [ [izi - ] loo] nda

                      H
                      |\
Delinking           [ [izi - ] loo] nda

                        H
                        |   \
Depressor Shift     [ [izi - ] loo] nda

Low Default         [ [ìzì- ] lóò] ndà

S.R.                  [izi - lóònda]
```

In (28) we observe that the Depressor Shift rule, unlike H tone spreading, is not blocked by a tone domain boundary. It is a special kind of local shift which applies in the word domain. Note that the Depressor Shift is not a domain-limit rule. Thus in (28) the shifted H tone does not affect the rightmost mora in the WD.

Lastly, I argue that H nouns such as *ín-kôsi* 'chief' (bi-syllabic) and *ú-bhóbhóyi* 'African hoopoe' (tri-syllabic) have a multiply linked underlying H tone on the stem (29).

(29) a. *ín-kôsi* 'chief'
H H
| /\
in - kosi

b. *ú-bhóbhóyi* 'African hoopoe'
H H
| / \
u - bhobhoyi

As seen in *ín-kôsi* the H tone does not surface on the last syllable. Nonetheless I argue that the underlying tonal pattern proposed in (19) exists, and that during the derivation process this pattern gets modified by some phonological processes.

At phrase medial position *ín-kôsi* surfaces with a H tone on the last syllable (30).

(30) *n-kosí yám* 'my chief/my God'

Compare (30) with (31). In (31) the noun *í-háshe* 'horse' has no underlying H tone on the first syllable, hence it has no H tone on the final syllable at phrase medial position.

(31) *háshe lám* 'my horse'

Because of its surface tonal structure at phrase medial position, I argue that *ín-kôsi* has an underlying H tone that is multiply linked to both syllables of the stem and that at phrase medial position the H tone is delinked from the penult by the general rule of Left Branch Delinking. The delinking of the left branch at phrase medial position would be made possible by the fact that in this position the penultimate vowel does not lengthen. I therefore assume that in isolation the lengthening of the penultimate vowel induces the fission of the multiply linked H tone. This then yields a situation which can account for the presence of the falling tone on the penult (32).

(32) U.R. H H
| /\
in-kosi

D-Structure H H

```
                     |        /\
                   [ [ [in- ]  [kosi] ] ]

Lengthening          H          H
                     |        /    \
                   [ [ [in- ]  [koo ]  [si] ] ]

H Fission            H        H  H
                     |        |   |
                   [ [ [in- ]  [koosi] ] ]

Spreading            N/A
```

In (32) we can see that after H Fission the H tone linked to the penult cannot spread to the second mora since there is a H tone on the final syllable. Now the only problem for the analysis proposed in (32) is the presence of the H tone on the final syllable. To resolve this problem I propose a rule which deletes the final H if there is an adjacent H (33). I call this rule Final H Deletion.

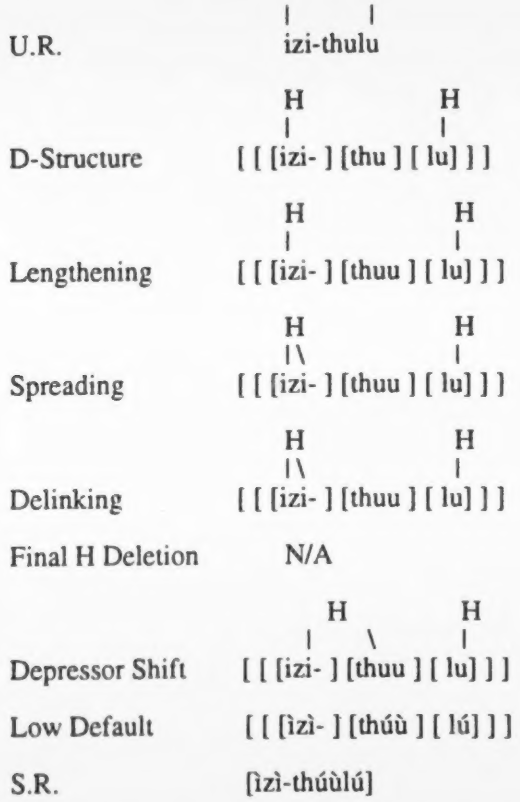
```
(33)  Final H Deletion        H          H        H------> Ø
                              |          |        |
                          [ [ [in- ]   [koo ]   [si] ] ]

      Low Default         [ [ [ín- ]   [kóò ]   [sì] ] ]

      S.R.                    [ín-kóòsi]
```

The Final H Deletion is ordered after the H tone spread, otherwise the H tone spread will apply and yield incorrect results (34).

```
(34)  Final H Deletion    H        H        H ------> Ø
                          |        |        |
                      [ [ [in- ]  [koo ]  [si] ] ]

      Spreading           H         H
                          |         |\
                      [ [ [in- ]  [koo ]  [si] ] ]

      Low Default     [ [ [ín- ]  [kóó ]  [sì] ] ]

      S.R.                * [ín-kóósi]
```

When taking into account the tonal structure of nouns such as *izi-thûlú*, 'deafs' where there is a H tone on the penult and a H tone on the final syllable, it would seem as if the rule of Final H Deletion lacks independent motivation. However, underlyingly the stem of *izi-thûlú*, has a H tone only on the final syllable as seen in '*Akú kho zi-thulú*, i.e. 'There are no deafs'. Thus the falling tone on the penult is the product of the Depressor Shift discussed above. It has the derivation given in (35).

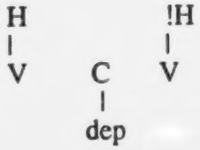
```
(35)                          H        H
```


U.R.
izi-thulu
D-Structure
H H
[ [ [izi- ] [thu ] [ lu] ] ]
Lengthening
H H
[ [ [izi- ] [thuu ] [ lu] ] ]
Spreading
H H
[ [ [izi- ] [thuu ] [ lu] ] ]
Delinking
H H
[ [ [izi- ] [thuu ] [ lu] ] ]
Final H Deletion
N/A
Depressor Shift
H H
[ [ [izi- ] [thuu ] [ lu] ] ]
Low Default
[ [ [ìzì- ] [thúù ] [ lú] ] ]
S.R.
[ìzì-thúùlú]

From the derivation in (35) we notice that the application of the depressor shift rule shifts the H tone onto the penult, thus yielding a situation which would necessitate the application of the Final H Deletion. However, the Final H Deletion is ordered before Depressor Shift and thus does not apply to the output of that rule.

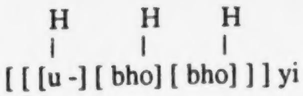
It is a generally assumed that depressor consonants have a L tone associated with them (Khumalo 1989). This assumption stems from the fact that when H tones are separated by a depressor consonant the second H tone will be downstepped (36).

(36)



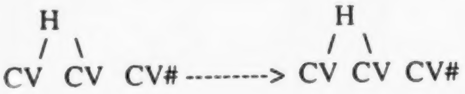
The downstep shown in (36) is attributed to the presence of the L tone on the depressor consonant. This phenomenon is noticed in *ú-bhóbhóyi*. I therefore assume that in *ú-bhóbhóyi* the depressor consonant *-bh-*, like vowel lengthening, induces fission of a multiply linked H tone, thus yielding the structure given in (37).

(37)



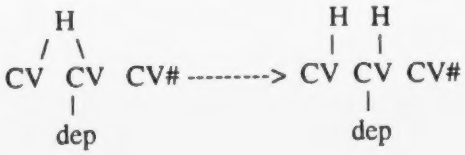
It is important to note that although the pattern HHO# is shown to be attested, it is limited to nouns that have a depressor consonant on the penult. I thus argue that the patterns OHO# and HHO# get neutralized by the Left Branch Delinking rule which leaves one pattern: OHO#. (See the neutralization schema in 38)). The neutralization process explains why the pattern HHO# surfaces only in nouns with a depressor consonant on the penult.

(38)



but

(38)



## 4. Conclusion.

I have shown that the tone domain is triggered by a H tone. This is the reason why L nouns do not have a sequence of tone domains despite the fact that they have the same morphological structure with H nouns.

The domains provide a solution to the problems that cannot be resolved by tonal analysis. For example, the tonal analysis does not show why certain syllables are targets for the H spread rule while others are not. Secondly, the spreading of a H tone cannot be explained adequately without postulating two rules of H spread. The domain-based approach, on the other hand, has shown that there is one H spread rule and that it spreads the H tone to the right edge of a tone domain.

I have shown that the H spread rule interacts with other rules to derive the surface tonal structure. In most cases the interaction of these processes involves the ordering of rules. The crucial rules that need to be ordered are: Vowel Lengthening before H Spreading, F H Deletion before Depressor Shift, and H Spread before F H Deletion.

Finally, I have shown that the complex tonal structure of Xhosa nouns can be explained by employing the theory of domains. The success of the rules in deriving correct surface forms while observing general principles such as the OCP gives them a valid status in the theory of phonology. In this way Xhosa renders support to the developing theory of domains.

# REFERENCES

Cassimjeee, F. and C.W. Kisseberth. 1992. Shingazidja metrical structure. To appear in the Proceedings of the 28th Regional Meeting of the Chicago Linguistic Society.

Clark, M. 1988. An accentual analysis of the Zulu noun: in Hulst, H and H. Smith (eds). Autosegmental studies on pitch accent. Dordrecht: Foris Publications.

Cole, J. and J. Coleman. 1992. No need for cyclicity in generative phonology. To appear in the Proceedings of the 28th Regional Meeting of the Chicago Linguistic Society.

Hyman, L. 1990. Boundary tonology and the prosodic hierarchy: in Inkelas, S. and D. Zec (eds.). The phonology-syntax connection. The University of Chicago Press: Chicago.

Khumalo, J.S. 1989. Zulu tonology. PhD Dissertation, Witwatersrand University.

Kisseberth, C.W. 1992. On domains. To appear in the Proceedings of the 28th Regional Meeting of the Chicago Linguistic Society.

McLaren, J. 1963. A new concise Xhosa-English dictionary. Maskew Miller Longman: Cape Town.

Pahl, H. W. 1976. IsiXhosa. Educum Publishers: Johannesburg.

# ON THE REPRESENTATION OF MANDARIN SYLLABLE STRUCTURE

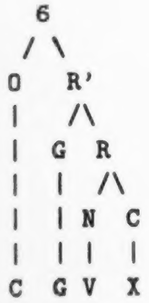
**Feng-Lan Kuo**
**University of Illinois**

Abstract: Duanmu (1990) argues that all Chinese dialects have a uniform syllabic structure of CVC. In his proposal, the pre-nucleus glide is part of the onset, where CG is a complex segment $C^G$. In this paper, I propose that Mandarin has a fixed syllabic template of CGVX, with one slot in the onset and three slots in the rimeprime (as projection of the rime). I claim that the pre-nucleus glide is obligatory, is an independent constituent, and that the pre-nucleus glide is adjoined to the rime constituent. Extensive evidence for this proposed template will be drawn from the analysis of derived mid vowels, distributional constraints on syllable structure, reduplication in language games, and rhyming in contemporary poetry and folksong.

## 1 Introduction

Mandarin is a language of monosyllabic morphemes. Studies of its syllable structure are seen in R. Cheng (1966), Fudge (1968), Chao (1968), C. Cheng (1973), Lin (1989), Bao (1990) and Duanmu (1990).[1] In this paper, I argue that Mandarin has a fixed syllabic structure of four slots, one in the onset and three in the rimeprime, as seen in (1).

(1) The Fixed Syllabic Structure of Mandarin



6 = syllable
O = onset
R', R = rimeprime, rime
G = glide
N = nucleus
C = coda

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 19-52

I argue that the fixed syllabic skeleton triggers three processes.[2] First, if the coda is not filled, the nuclear segment will spread to the coda, creating a long vowel. Hence, vowel length is predictable in Mandarin. Second, if the onset is not filled, then the following conventions are available to satisfy the obligatory onset requirement: (i) spread the nuclear segment (if [+high] or syllabic nasal) to the onset, (ii) spread the pre-nucleus glide to the onset, (iii) associate the pre-nucleus glide onto onset position, or (iv) the onset is specified as what is commonly called the 'zero onset', which I represent by the symbol #. Third, if the pre-nucleus glide position is not filled, then the glide position is specified as $\phi$, a null segment, by default. The analysis of these three processes follows from the analysis of fixed syllable structure in Mandarin, as argued in the following sections. Specifically, I claim that in Mandarin:

(2) a. Every rime has two X slots.
b. Every syllable has an obligatory onset.
c. Every syllable has an obligatory pre-nucleus glide.

Evidence for (2) comes from duration of the rime, suffixation processes, reduplication, and distributional constraints on syllable structure.

The phonetic inventories in Mandarin are as follows.

(3) Consonant Inventory:

| | | | | |
|---|---|---|---|---|
| p | p$^h$ | m | f | |
| t | t$^h$ | n | | l |
| ts | ts$^h$ | | s | |
| tṣ | tṣ$^h$ | | ṣ | ṛ |
| tς | tς$^h$ | | ς | |
| k | k$^h$ | | x | (#) |

(4) Vowel Inventory: (6 = schwa)

| | | | |
|---|---|---|---|
| i | ü | I | u |
| e | | 6 | o |
| | | a | |

**The Rime has two X Slots:** The fact that no segment can follow a VG sequence within any syllable in Mandarin shows that there are at most two weight-bearing slots in the rime. In addition, when the diminutive suffix /r/

is added to a syllable, as in (5), it replaces the original coda, instead of concatenating to the original syllable, also suggests that every Mandarin syllable has a fixed weight of two slots in the rime (cf. Chao (1968), Duanmu (1990)).

(5) Mandarin Diminutive Suffixation
ya + r ⟶ yar
yan + r ⟶ yar (*yanr)
gwa + r ⟶ gwar
gway + r ⟶ gwar (*gwayr)

A third argument for the fixed rime weight is found in Howie (1976) and Duanmu's (1990) studies; all regular Mandarin syllables have similar duration, which indicates that the rime has a fixed weight.[3] For example, the fact that /ta/, /tan/, /tau/, and /tiau/ have similar duration suggests that the vowel in an open syllable is long. Duanmu (1990) shows that vowel length is predictable, and that the long vowel in an open syllable is the result of spreading to the vacant 2nd rime position. The fact that the existence of the pre-nucleus glide does not affect the duration of the whole syllable when followed by VX sequence, along with the pattern seen in (5), shows that the pre-nucleus glide is not a weight-bearing unit.

**Obligatory 'Zero Onset'**: According to Chao (1968), and Duanmu (1990), Mandarin syllables that are not written with an onset have a 'zero onset'. If the syllable nucleus is a high vowel [i,u,ü], or the syllabic nasal [m], then the zero onset is [y w ÿ m] respectively. However, I make a stronger claim, that syllable-initial glides also can spread onto onset position, or can be mapped onto onset position. This will be seen in the evidence from reduplication in section 3. If there are no [+high] segments in the nucleus or pre-nucleus position, or no syllabic nasals in the nucleus position, then the zero onset has the following four variants:[4]

(6) a. velar nasal [ng]
b. velar or uvular unaspirated fricative/continuant [$\tau$]
c. glottal stop [?]
d. glottal unaspirated continuant [H] (which Chao calls a 'true vowel' onset).

Apart from the questions of why these four variants are selected, and what the relationship between these variants is, I agree with Duanmu that the zero onset is not a phonetically motivated phenomenon as is the case in English. Rather it is a phonologically motivated phenomenon. Its presence prevents the resyllabification of the [-high] nucleus vowel with the preceding coda in pronunciation as shown in (7).

(7) /mian ao/ ---> a. [myan ?aw] (Duanmu (1990:20))
'cotton coat' b. [myan ɣaw]
c. [myang ɣaw]
d. [myang ngaw]
e.*[mya.naw]

In this paper, I argue that the zero onset phenomenon does not result from an obligatory condition, but rather results from a principle with fixed parametric variations across dialects. I term this the Onset Satisfaction Principle (hereafter OSP). The evidence for this principle comes from the observation of the behavior of the pre-nucleus glide in GV syllables in various language games, and in rhyming. The Onset Satisfaction Principle adopted here is stated in (8). The details of this principle, as it applies in various language games, are considered below in section 3.

(8) Three Parameters of Onset Satisfaction Principle
a. Spread [+high] from the nucleus or pre-nucleus segment onto the onset position, otherwise specify the zero onset #.
b. Insert the zero onset # in every vowel initial syllable.
c. Associate the front glide [y] onto onset position in GV syllables.

The Mandarin-based language games: May-ka, Mey-ka, and Man-t'a are languages which utilize clause (a) of the OSP to satisfy the onset requirement.[5] Taiwanese and Taiwanese-based language games are instances of languages which make use of clause (b) of the OSP to meet the obligatory onset requirement.[6] An example which uses clause (c) of the OSP can be found in Mo-pa, a Kushan-based language game.[7]

**Obligatory 'Zero Glide'**: In this paper I argue that every syllable in Mandarin has an obligatory glide, and that the 'zero glide' phenomenon is phonologically motivated.[8] Specifically, its presence prevents violations of the Labial Constraint in syllables such as /bu/ and /pu/, as discussed in section 2. The zero glide also helps explain a seeming asymmetry between the front glides [y,ỹ] and the back glide [w] in reduplication and language games, as discussed in section 3. In addition, it explains why rhyming between [aw] and [yaw], and between [ỹan] and [an] are possible, as will be discussed in section 5. I propose that the pre-nucleus glide position can be filled only by [+high] segments or by the spreading of [+high] segments; in the absence of a [+high] segment, the default $\phi$, a null segment, is inserted to fill the syllable template.

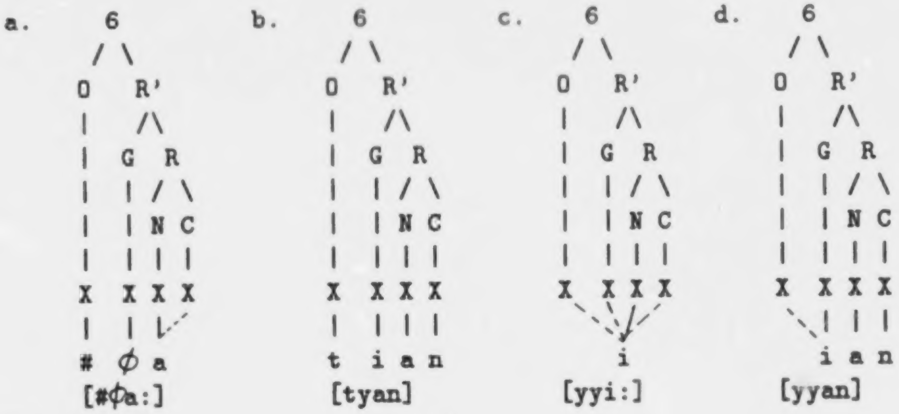
Given a string, along with my proposed syllable template, I assume the following syllabification (template mapping) algorithm:

(9) a. Link the most sonorant segment to the nucleus.
b. Link the onset if there is [+cons] segment available.
c. Link the coda.
d. Link the pre-nucleus glide.
e. If the coda is not filled, spread the nuclear segment to it.
f. If the onset is not filled, consult the OSP to insert a segment into it.
g. If the pre-nucleus glide position is empty, insert the default zero glide $\phi$.

For the following 4 representative syllable types in Mandarin, my proposed syllabification algorithm will assign the syllable representations as shown in (11).

(10) a. /a/ 'filthy' b. /tian/ 'heaven'
c. /i/ 'clothes' d. /ian/ 'smoke'

(11)



(# = zero onset)

More arguments for the status of the pre-nucleus glide and my proposed syllable structure will be drawn from the analysis of derived mid vowels, distributional constraints on syllable structure, reduplication and language games, and rhyming in poetry and folksongs.

**Analysis of Derived Mid Vowels**: In Mandarin, the mid vowel /6/ acquires its [back] (and [round]) specification from an adjacent glide on either side. If glides are present on both sides, then it assimilates to the segment on the right (the coda) (cf. C. Cheng (1973)).

(12) Distribution of Mid Vowels
  a. /b6i/ ⟶ [bey] 'cup'  /bi6/ ⟶ [bye] 'separate'
  b. /t6u/ ⟶ [tow] 'steal'  /tu6/ ⟶ [two] 'delay'
  c. /di6u/ ⟶ [dyow] 'throw'  /tu6i/ ⟶ [twey] 'withdraw'

With the syllable structure in (1), I can account for the pattern of assimilation by stating that the mid vowel assimilates to the structurally closest segment. Under this formulation, directionality of assimilation needs not be stipulated.

I turn next to the distributional constraints on labial and back segments in section 2. Section 3 sketches reduplication patterns seen in the various Fanqie languages: Mo-pa, May-ka, Mey-ka, Na-ma, and Taiwanese. Section 4 points to several problems in Duanmu's (1990) and Bao's (1990) analyses. Sections 5 demonstrates how rhyming in poetry and folksongs, provides some insight to the status of the pre-nucleus glide and the syllable structure of Mandarin. Some residual problems are addressed in the concluding section.

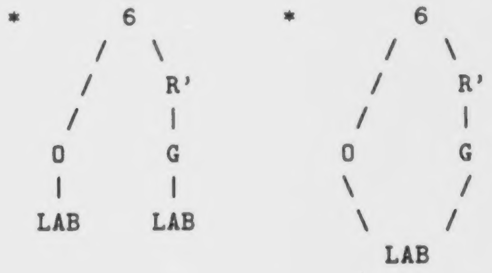
## 2 Distributional Constraints on Syllable Structure

If the pre-nucleus glide is part of the onset, I would expect there to be some co-occurrence restrictions for consonant clusters in the onset position. However, I find almost all kinds of consonants freely occurring before the front glide [y]. Before the back glide [w], only labial consonants are prohibited. The data in (13) illustrate the relevant distribution of the consonants before the two glides. I argue that labial consonants are disallowed before the back glide [w] due to the Labial Constraint. My proposed Labial Constraint is given in (14).

(13) Data Showing Distribution of Labial Segments:

| | | | |
|---|---|---|---|
| a. labials | | | |
| byan | 'change' | pyan | 'cheating' |
| myan | 'noodle' | *fyan[9] | |
| *bwan | *pwan | *mwan | *fwan |
| b. dentals | | | |
| tyan | 'heaven' | dyan | 'shop' |
| nyan | 'read' | lyan | 'in love' |
| twan | 'fast stream' | dwan | 'hold' |
| nwan | 'warm' | lwan | 'twin' |
| c. palatals | | | |
| jyan | 'sharp' | qyan | 'money' |
| xyan | 'thread'[10] | | |
| d. velars | | | |
| gwan | 'close' | kwan | 'wide' |
| hwan | 'happy' | | |

(14) The Labial Constraint



The Labial Constraint says that the onset and the pre-nucleus glide of a syllable can not both be labials, no matter what kind of linking to the Labial node is. The Labial Constraint allows the occurrence of the following labial sequences.

a. bu pu mu fu
(15) b. baw paw maw fow

The Labial Constraint directly explains why the forms in (15b) are good, and why *bwan, *pwan, etc. are bad. As for the forms in (15a), I argue that the presence of the zero glide $\phi$ prevents these forms from violating the Labial Constraint. Therefore, the forms in (15a) are well-formed. In my analysis, /bu/ and */bua/ have the following syllabic representations:

```
(16) a.    6            b.*    6
          / \                 / \
         O   R'              O   R'
         |   /\              |   /\
         |  G  R             |  G  R
         |  | / \            |  | / \
         X  X X X            X  X X X
         |  | | /            |  | | /
         b  ϕ u              b  u a
```

In addition to the Labial Constraint, which holds between the onset and the pre-nucleus glide due to **string adjacency**, I find that there is a co-occurrence restriction holding between the pre-nucleus glide and the rime. If there are three elements within the rime constituent (R'), then the pre-nucleus glide and the post-nucleus glide must have opposite specifications for backness (cf. C. Cheng 1973). In this paper, I give a formal representation of the back dissimilation constraint described by Cheng, given in (17). I argue that the co-occurrence restriction between the pre-nucleus glide and the post-nucleus glide is better expressed as a co-occurrence restriction between the pre-nucleus glide and the rime (R), due to **structural adjacency**.

(17) The Back Dissimilation Condition

```
            R'
          /   \
         G     R
         |     |
   [α back] [-α back]
```

The relevant data for the ill-formed patterns *yVy and *wVw are shown in (18).

(18) Data Showing Back Constraint:

a. *bwaw *pwaw *mwaw *fwaw
b. *dwaw *twaw *gwaw *kwaw
c. *bwow *pwow *mwow *fwow
d. *dwow *twow *gwow *kwow
e. *byay *pyay *myay *fyay
f. *dyay *tyay *gyay *kyay
g. *byey *pyey *myey *fyey
h. *dyey *tyey *gyey *kyey[11]

These examples motivate my proposed Back Dissimilation Condition. Here I assume that the low vowel /a/ and the mid vowel /6/ are not specified for the feature [back]. Recall that the mid vowels [e] and [o] are derived from /6/. Therefore, the redundancy rules supplying the feature [+back] must apply after the Back Dissimilation Condition applies. The matrix in (19) shows the underspecified representation of the vowels in Mandarin.

(19) Distinctive Feature Specification: (6 = schwa)

| | i | u | u | 6 | a |
|---|---|---|---|---|---|
| back | - | - | + | | |
| round | | + | + | | |
| high | + | + | + | | |
| low | | | | | + |

Redundancy Rules:
[ ] --> [+back]
[ ] --> [-round]
[ ] --> [-high]
[ ] --> [-low]

The advantage of the Back Dissimilation Condition is that I can account for the non-occurrence of the patterns *wVw, and *yVy by a single constraint. Also, the benefit of my proposed obligatory zero glide is that it accounts for the previously unexplained exception of the lexical item [yay] 'cliff' to the Back Dissimilation Condition (cf. Hockett (1947), Fudge (1968), Cheng (1973), and Fu (1990)). I argue that the representation for the lexical item 'cliff' is [yφay]. Therefore there is no violation of the Back Dissimilation Condition.

It is interesting to find that the Back Dissimilation Condition functions in another Chinese dialect: Kejia (cf. Chung (1988)). The Back Dissimilation Condition predicts that the combinations in (20) are all ungrammatical in Kejia.[12] Notice that the mid vowels in Kejia appear in underlying vowel inventories, and so are specified for the feature [back].

| | | |
|---|---|---|
| | *yey | *waw |
| (20) | *yay | *wow |
| | *ye | *wo |

Further support for the obligatory zero glide, the Back Dissimilation Condition and the Labial Constraint will be seen in section 3, where reduplication and language games are discussed. It seems to me that the Back Dissimilation Condition is not an accident; rather it is a systematic restriction on the syllable structure of languages such as Mandarin, Kejia, and Fanqie languages.

## 3 Language Games and Reduplication

In this section, I bring in evidence from language games, known as Fanqie languages, to provide insight to the status of the pre-nucleus glide and the syllable structure of Mandarin (and possibly other Chinese dialects as well). Here I adopt Steriade's (1988) model of reduplication. She argues that reduplication is a process of total copying of the base, followed by insertion or truncation, which operates on the string derived through the total copying of the base. The various Fanqie languages are formed by a process of full reduplication, followed by language-specific conventions of inserting some segmental material into some prosodic constituent(s).

There are three possible ways to incorporate the pre-nucleus glide into syllable structure: (i) make it part of the onset, (ii) make it part of the rime, and (iii) make it an independent constituent. I argue that the facts of reduplication can be used to determine the syllable constituency of the pre-nucleus glide. If insertion into the onset causes the substitution of the pre-nucleus glide, the simplest analysis would take the pre-nucleus glide to be part of the onset. If insertion into the rime causes substitution of the pre-nucleus glide, the simplest analysis would take the pre-nucleus glide to be part of the rime. If neither onset insertion, nor rime insertion cause substitution of the pre-nucleus glide, then the simplest analysis would take the pre-nucleus glide to be an independent constituent.

The differing status of the pre-nucleus glide can be seen from three representative Fanqie languages of Chinese: Mo-pa, May-ka, and Na-ma. In

Mo-pa, the pre-nucleus glide behaves as part of the rime. In May-ka, it manifests itself as an independent constituent. In Na-ma, it functions like part of the onset. I argue that, in the various language games, the pre-nucleus glide always appears as an independent constituent, which is adjoined to the rime.

My solution to the reduplication patterns seen in these three Fanqie languages will mainly rely on the independent status of the pre-nucleus glide, the richer syllable structure, and the Onset Satisfaction Principle to derive the correct output.

**Mo-pa**: Mo-pa is a language game based on the Kunshan dialect. When insertion operates on the rime of the first syllable, it causes the substitution of the pre-nucleus glide. This suggests that the pre-nucleus glide is part of the rime. I suggest the rules as shown in (21) for Mo-pa and a sample derivation is given in (22) for illustration.

(21) a. Reduplicate the syllable.
b. Insert [ϕo] into the first rime (R').
c. Switch the value of [cont] of the second onset.

(22) lyã
lyã-lyã (21a)
lϕo-lyã (21b)
lϕo-tyã (21c)
output: lϕo-tyã[lo-tyã]

Consider an instance of a GV syllable from Mo-pa. My analysis for the example /iO/ 'want' from Chao (1931) is given in (23).

(23) iO
yϕO (OSP, clause (c): associate [y] onto onset position)
yϕO-yϕO (21a)
yϕo-yϕO (21b)
yϕo-tçϕO (21c)
output: yϕo-tçϕO[yo-tçO]

In Mo-pa, I find that the front glide [y] in a GV sequence is mapped onto onset position, due to clause (c) of the OSP. Unfortunate ly, Chao (1931) does not include any data beginning with the back glide [w] in a GV sequence. Therefore it is hard to know whether [w] in a GV sequence in Mo-pa is mapped onto the onset position. However, the OSP gives me the correct output for the above example. As we go on, we will find that the OSP is a principle with fixed parametric variations across dialects.

**May-ka**: May-ka is a language game based on Mandarin. When insertion operates on the rime of the first syllable and the onset of the second syllable, it does not cause the substitution of the pre-nucleus glide in either syllable. This indicates that the pre-nucleus glide has an independent status in the syllable. I suggest the following rules and derivations for May-ka.

(24) a. Reduplicate the syllable.
b. Insert [ay] into the first rime (R).
c. Insert [k] into the second onset.

(25) hwey
hwey-hwey (24a)
hway-hwey (24b)
hway-kwey (24c)
output: hway-kwey

(26) lya
lya-lya (24a)
lyay-lya (24b)
lyay-kya (24c)
lye-tçya (Repair and Palatalization)[13]
output: lye-tçya

Consider two examples of GV syllables from May-ka, one beginning with the front glide [y], and the other with the back glide [w]. For the example /iang/ 'sun', I propose the analysis in (27). Again, it is crucial to make reference to the OSP in my analysis.

(27) iang
yyang (OSP, clause (a): spreading)
yyang-yyang (24a)
yyay-yyang (24b)
yyay-kyang (24c)
yye-tçyang (Repair and Palatalization)
output: yye-tçyang[ye-tçyang]

Notice that after the application of (24c), there is a process I term Repair involved in the derivation in (26) and (27). I suggest the Repair rule: ay ⟶ e / y__ for May-ka, where [y] is in the pre-nucleus position. I argue that the motivation for this rule is due to the Back Dissimilation Condition proposed earlier in this paper. I find that this constraint is not only respected in Mandarin, but also in the language game May-ka.

In accounting for GV syllables beginning with the back glide [w], I need to consider an additional Mandarin rule which changes [w] in syllable initial position to [v].[14] In Duanmu (1990), this rule is taken to be obligatory and is carried over to May-ka. However, in order to account for the data in Chao (1931), I must follow Yip (1982) and Bao (1990) in assuming that this rule is optional. This allows the following two analyses for the example /uan/ 'curve'.

(28) uan
wwan (OSP: spreading)
w$\phi$an (Repair: due to violation of the Labial Constraint)[15]
v$\phi$an (rule w → v)
v$\phi$an-v$\phi$an (24a)
v$\phi$ay-v$\phi$an (24b)
v$\phi$ay-k$\phi$an (24c)
output: v$\phi$ay-k$\phi$an[vay-kan]

(29) uan
wwan (OSP: spreading)
w$\phi$an (Repair: due to violation of the Labial Constraint)
w$\phi$an-w$\phi$an (24a)
w$\phi$ay-w$\phi$an (24b)
w$\phi$ay-k$\phi$an (24c)
output: w$\phi$ay-k$\phi$an[way-kan]

Notice that in (28) and (29), there is another Repair rule involved in the derivations. I suggest the Repair rule: w ⟶ $\phi$ / w __ for Mandarin and May-ka. I argue that the motivation for this Repair rule is due to the Labial Constraint proposed earlier in this paper. In proposing this additional Repair rule, my analysis explains a seeming asymmetry between the front glide and the back glide, where the front glide [y] surfaces in the pre-nucleus position of the second syllable, but the back [w] does not. My analysis requires invoking the Onset Satisfaction Principle and the Labial Constraint to account for this type of data. I argue that the pre-nucleus glide can be a potential onset through spreading. May-ka, like Mandarin, spreads [+high] segments to the onset position to meet the requirement of the Onset Satisfaction Principle. This analysis accounts for both kinds of output.

I turn now to Mey-ka, another Fanqie language based on Mandarin, which shows similar phonological patterns to May-ka.

**Mey-ka**: Mey-ka is another Mandarin-based language game, which provides further support for my proposed Back Dissimilation Condition and the Onset Satisfaction Principle. I suggest the following rules for Mey-ka and some derivations are seen in (31) through (33).

(30) a. Reduplicate the syllable.
b. Insert [ey] into the first rime (R).
c. Insert [k] into the second onset.

(31) lya
lya-lya (30a)
lyey-lya (30b)
lyey-kya (30c)
lϕey-kya (Repair: due to violation of the Back Dissimilation Condition)[16]
output: lϕey-kya[ley-kya]

(32) iang
yyang (OSP: spreading)
yyang-yyang (30a)
yyey-yyang (30b)
yyey-kyang (30c)
yϕey-kyang (Repair)
output: yϕey-kyang[yey-kyang][17]

(33) hwey
hwey-hwey (30a)
hwey-hwey (30b)
hwey-kwey (30c)
output: hwey-kwey

In comparing (25) through (29) with (31) through (33), I find that although May-ka and Mey-ka are both based on the same source language, Mandarin, they differ in the strategies that they employ to repair the ill-formed syllables that are produced by reduplication and the insertion of segmental material. In May-ka, [ay] is changed to [e] when preceded by the

the front glide [y], whereas in Mey-ka, the front glide [y] gets deleted. I suggest the Repair rule: y $\longrightarrow \phi$ / __ Vy for Mey-ka. The motivation for the [y] deletion in Mey-ka, seen in (31) and (32), again comes from the Back Dissimilation Condition. Again, I utilize the OSP to derive correct output for the glide-initial syllable in Mey-ka. Thus the output form in (32) is not an exception to the Back Dissimilation Condition. Rather, the presence of the zero glide causes the output to conform to the Back Dissimilation Condition.

**Na-ma**: In the Na-ma case, after the insertion of segmental material into the first onset, the pre-nucleus glide is lost in the output, which suggests that the glide is part of the onset. Such cases provide a challenge to any analysis which assumes that the pre-nucleus glide is part of the rime. In this section, I demonstrate that my analysis is able to handle this phenomenon. The rules for Na-ma are shown in (34) and a derivation is given in (35).

(34) a. Reduplicate the syllable.
b. Template for the first syllable of the redupicant: CVX
c. Insert [n] into the first onset.

(35) twey
twey-twey (34a)
tey-twey (34b)
ney-twey (34c)
output: ney-twey

The formulation of (34b) suggests prosodic circumscription operates on the first syllable of the reduplicant.

**Taiwanese**: To complete my proposed Onset Satisfaction Principle, I introduce another Fanqie language, which is based on Taiwanese. According to the description of Li (1985), all vocalic-initial syllables in Taiwanese have a zero onset #, which has only one phonetic alternant; i.e., the glottal stop [ʔ]. The fact that when insertion operates on the onset, the pre-nucleus glide remains intact suggests that the glide belongs to the rime. The rules and examples for the Taiwanese-based language game can be seen in (36) through (38).

(36) a. Reduplicate the syllable.
b. Insert [l] into the first onset.
c. Insert [i] into the second nucleus (or rime).[18]

(37) Taiwanese-based Language Games:

| | (36a) | | (36b) | | (36c) | |
|---|---|---|---|---|---|---|
| kun | ⟶ | kun-kun | ⟶ | lun-kun | ⟶ | lun-kin |
| sat | ⟶ | sat-sat | ⟶ | lat-sat | ⟶ | lat-sit |
| a | ⟶ | a-a | ⟶ | la-a | ⟶ | la-i |
| e | ⟶ | e-e | ⟶ | le-e | ⟶ | le-i |
| ȳā | ⟶ | ȳā-ȳā | ⟶ | nȳā-ȳā | ⟶ | nȳā-ī[19] |
| hwe | ⟶ | hwe-hwe | ⟶ | lwe-hwe | ⟶ | lwe-hi |
| tsya | ⟶ | tsya-tsya | ⟶ | lya-tsya | ⟶ | lya-tsi |
| tsay | ⟶ | tsay-tsay | ⟶ | lay-tsay | ⟶ | lay-tsi |
| t$^h$aw | ⟶ | t$^h$aw-t$^h$aw | ⟶ | law-t$^h$aw | ⟶ | law-t$^h$i |
| tsyaw | ⟶ | tsyaw-tsyaw | ⟶ | lyaw-tsyaw | ⟶ | lyaw-tsi |
| k$^h$way | ⟶ | k$^h$way-k$^h$way | ⟶ | lway-k$^h$way | ⟶ | lway-k$^h$i[20] |

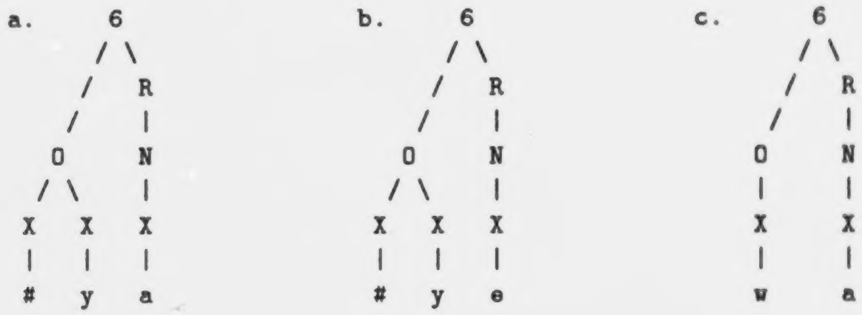
In summary, I have shown that the Onset Satisfaction Principle plays an important role in accounting for glide-initial words in Fanqie languages. My analysis suggests that all the Fanqie languages employ the same Fanqie rules: full reduplication and constituent modification. However, they may differ in the parametric invoking of the OSP, and in which constituent is chosen for insertion of segmental material. Data from various Fanqie languages investigated here support my claim that the pre-nucleus glide is a part of the rime. The benefit of my analysis is that there is just one single syllable structure required for both the base language and its corresponding Fanqie language(s).

## 4 Previous Analyses

In this section I present the analyses proposed by Bao (1990) and Duanmu (1990), and then discuss the differences between their and my analyses.

**Bao's Analysis:** Bao (1990) argues that different Chinese dialects can have different syllable structures, and that there is not a fixed template for syllable structure within a single dialect.[21] He proposes that the pre-nucleus is part of the onset, where the CG sequence is a consonant cluster in Mandarin and Mandarin-based Fanqie languages. In his analysis, different pre-nucleus glides have different syllable status when they occur in syllable-initial position. Therefore, /ya/ 'crow', /ÿ6/ 'moon', and /wa/ 'frog' have the following syllable representations.

(38) Bao's Syllable Representations: (Bao 1990:334)



Based on these assumptions, he can account for the following patterns shown in the language games of May-ka and Mey-ka.

(39) May-ka
wan
wan-wan (Reduplication)
way-wan (Replace the first rime with [ay])
way-kan (Replace the second onset-initial with [k])
output: way-kan

(40) May-ka
lya
lya-lya (Reduplication)
lyay-lya (Replace the first rime with [ay])
lyay-kya (Replace the second onset-initial with [k])
lye-kya (Repair and Palatalization)
output: lye-tçya

(41) May-ka
yang
#yang
#yang-#yang (Reduplication)
#yay-#yang (Replace the first rime with [ay])
#yay-kyang (Replace the second onset-initial with [k])
#ye-tçyang (Repair and Palatalization)
output: #ye-tçyang[22]

(42) Mey-ka lya
lya-lya (Reduplication)
lyey-lya (Replace the first rime with [ey])
lyey-kya (Replace the second onset-initial with [k])
ley-kya (Repair)
output: ley-kya

(43) Mey-ka #yang
#yang-#yang (Reduplication)
#yey-#yang (Replace the first rime with [ey])
#yey-kyang (Replace second onset-initial with [k])
—————— (Repair: Not applied)
output: #yey-kyang

Bao proposes the Repair rule: $ay \rightarrow e\ /\ y\ \_\_$ for May-ka, and the Repair rule: $y \rightarrow \phi\ /\ \_\_\ ey$ for Mey-ka which are the same as the two Repair rules I propose in section 3. He argues that the application of the two Repair rules is due to the fact that Mandarin does not allow the form CyVy, and the same kind of constraint is carried over to the Mandarin-based language games: May-ka and Mey-ka. Though he does not spell out what kind of constraint it is, I find that his argument indicates that it is in essence comparable to my proposed Back Dissimilation Constraint. However, his analysis provides no explanation for why [#yay] 'cliff' is well-formed in Mandarin, why [#yey] in Mey-ka does not undergo the Repair rule, or why [#yay] in May-ka does undergo the Repair rule. It must be stipulated that the lexical item [#yay] is an exception to the back constraint, and it also must be stipulated that the front glide [y] is not deleted when preceded by the zero onset #. This raises the question of whether the zero onset # really behaves differently from a regular onset.

Bao also notes that the back glide [w] has several free phonetic variants, ranging from [w] to [v], when it occurs in syllable-initial position, whereas the front glide [y] always remains stable and preserves its glide status. Bao claims that these facts indicate that they have different phonological behaviors. In comparing Bao's analysis with my analysis seen in (26) through (29) and (31) through (32), my analysis differs from Bao's analysis in two ways: (i) my analysis accounts for the distribution of [back] segments in Mandarin, and the application of the two Repair rules in May-ka and Mey-ka in a principled way: by making reference to the OSP, the Back Dissimilation Constraint and the insertion of the zero glide; and (ii) with reference to the OSP, the Labial Constraint and the insertion of the zero glide, my analysis does not postulate any asymmetrical behavior between the two kinds of glides.

Another major problem for Bao's proposal is that his analysis allows substitution to operate on **part of a prosodic unit**, which is contrary to the spirit of Steriade's (1988) claim, in this case, the first member of the Onset. His analysis predicts that substitution can also operate on the second member of the Onset. However, such kind of case is never attested.

**Duanmu's Analysis:** Duanmu (1990) argues that all Chinese dialects have a uniform syllabic structure of three slots, one in the onset and two in the rime. He proposes that the pre-nucleus glide is part of the onset where CG is a complex segment $C^G$. In his analysis, both front glides and back glides in syllable-initial position are in the onset position.

In dealing with language games, since $C^G$ is a complex segment, Duanmu's analysis relies on the notion of what he termed Feature Recycling to get back the secondary articulation that was lost after substitution operates on the onset constituent.[23] However, his analysis has difficulties in accounting for the following data of May-ka. Note that the parenthesis indicates that the old segmental material still hangs around even after insertion of new segmental material.

| | | |
|---|---|---|
| (44) May-ka | wan | |
| | van | (rule w → v) |
| | van-van | (Reduplication) |
| | vay(an)-van | (Replace the first rime with [ay]) |
| | vay(an)-k(v)an | (Replace the second onset with [k]) |
| | vay-kan | (Feature Recycling) |
| | output: vay-kan[24] | |

In (45) I find that if the optional rule w → v /# ___ does not apply, then the incorrect output *[way-$k^w$an], instead of the correct output [way-kan], is derived in Duanmu's analysis.

| | | |
|---|---|---|
| (45) May-ka | wan | |
| | wan-wan | (Reduplication) |
| | way(an)-wan | (Replace the first rime with [ay]) |
| | way(an)-k(w)an | (Replace the second onset with [k]) |
| | way-$k^w$an | (Feature Recycling) |
| | output: *way-$k^w$an | |

In (46) and (47) we see that by not recognizing the back constraint as the driving force for the Repair rule, Duanmu's analysis fails.[25] The correct output for (46) should be [l$^y$e-tç$^y$a], whereas the correct output for (47) should be [ye-tç$^y$ang]. In comparing Duanmu's analysis with my analysis in (26) and (27), we find that the Back Dissimilation Constraint and the Repair rule are what we need, in order to account for this type of data.

(46) May-ka lya
l$^y$a-l$^y$a (Reduplication)
l$^y$ay(a)-l$^y$a (Replace the first rime with [ay])
l$^y$ay(a)-k(l$^y$)a (Replace the second onset with [k])
l$^y$ay-k$^y$a (Feature Recycling)
l$^y$ay-tç$^y$a (Palatalization??)[26]
output: *l$^y$ay-tç$^y$a

(47) May-ka yang
yang-yang (Reduplication)
yay(ang)-yang (Replace the first rime with [ay])
yay(ang)-k(y)ang (Replace the second onset with [k])
yay-k$^y$ang (Feature Recycling)
yay-tç$^y$ang (Palatalization??)
output: *yay-tç$^y$ang

In (48) through (50) we see examples of Duanmu's proposal applied to Mey-ka. Duanmu proposes that Onset Simplification is a mechanism that various language games can utilize to delete the minor articulator of a complex segment, in order to derive the correct output. He argues that Mo-pa is one of the languages that makes use of the Onset Simplification.[27] However, we find Onset Simplification can not be applied straightforwardly in Mey-ka. In (50), after the application of Onset Simplification, the incorrect form *[hey-k$^w$ey] is derived. One possible solution is to stipulate that Onset Simplification only targets the front glide. In this case, Duanmu provides no principled explanation for the asymmetry between the two glides. In comparing Duanmu's proposal applied to Mey-ka with the analysis I presented in (30) through (33), there is no need to appeal to the notion of Feature Recycling, nor the mechanism of Onset Simplification in my analysis.

(48) Mey-ka $l^y$a
$l^y$a-$l^y$a (Reduplication)
$l^y$ey(a)-$l^y$a (Replace the first rime with [ey])
$l^y$ey(a)-k($l^y$)a (Replace the second onset with [k])
$l^y$ey-$k^y$a (Feature Recycling)
ley-$k^y$a (Onset Simplification)[28]
output: ley-$k^y$a

(49) Mey-ka yang
yang-yang (Reduplication)
yey(ang)-yang (Replace the first rime with [ey])
yey(ang)-k(y)ang (Replace the second onset with [k])
yey-$k^y$ang (Feature Recycling)
———— (Onset Simplification: not applied)
output: yey-$k^y$ang

(50) Mey-ka $h^w$ey
$h^w$ey-$h^w$ey
$h^w$ey(ey)-$h^w$ey
$h^w$ey(ey)-k($h^w$)ey
$h^w$ey-$k^w$ey (Feature Recycling)
hey-$k^w$ey (Onset Simplification)
output: *hey-$k^w$ey

In summary, by appealing to my proposed OSP and the Labial Constraint, my analysis is able to account for the free variation between [vay-kan] and [way-kan] seen in (28) and (29). In addition, by recognizing the independent status of the pre-nucleus glide and the Back Dissimilation Condition, there is no need to appeal to the notion of Feature Recycling, nor to appeal to the mechanism of Onset Simplification. My analysis can account for all the data investigated here without any ad hoc stipulations. In addition, my analysis is able to maintain a symmetry in the phonological behavior of the glides.

## 5 Rhyming in Poetry and Folksongs

In this section, I will demonstrate that rhyming in poetry and folksongs reveals the status of the pre-nucleus glide, based on native speakers' identification of the rhymes. In Mandarin, two syllables rhyme if they share the same nucleus and coda. Wang (1973) claims that poetic rhyme does not include the pre-nucleus glide. I tested this hypothesis against native speakers' judgments in identifying rhymes, but found, constrary to Wang, that the pre-nucleus glide is part of the rime. Consider the following poem and folksong.

(51) yi ken tsi ts̲u ts̲i myaw **myaw**
song key baw baw ts̲wo kwan **ςyaw**
ςyaw er twey ts̲eng kow, kow er twey ts̲eng **ςyaw**
ςyaw zhong tς$^h$wey tς$^h$w ςin s̲in **tyaw**
ςyaw baw baw, yi ti yi ti ςw̃ey xwey lyaw, ςw̃ey xwey **lyaw**

'There is a piece of bamboo that is very straight.'
'Give it to the kid to make a flute.'
'Put the flute right in front of your mouth, put your mouth right in front of the flute.'
'Make a new popular song out of it.'
'Little kid learns to make the new song little by little, learn how to.'

In this folksong, the last word of each line rhymes. Among the twelve speakers tested, two of them characterized the rhyming constituent as [aw], four of them chose [yaw], and the rest recognized both [aw] and [yaw] as rhyming constituents. The traditional account of rhyming maintains that the pre-nucleus glide is outside the rhyming constituent. This account explains the first set of speakers. However, since for a large number of speakers, the pre-nucleus glide is a part of the rhyming constituent, it is necessary to identify a syllable constituent that includes the glide and the syllable rime. This constituent is R' in my proposed syllable structure. By identifying the constituent R' in addition to the constituent R, I can account for the first set of speakers by saying that for them, rhyming is scanning the R node (loose rhyme) alone. I am also able to account for the second set of speakers by saying that rhyming is scanning the R' node (strict rhyme), instead of the R node. As for the third set of speakers (most speakers), the free variation between [yaw] and [aw] follows naturally from my proposal that the pre-nucleus glide is an independent constituent, which is part of the rime. If the pre-nucleus glide belongs to the onset, then one has to explain how [yaw] can be a rhyming constituent. Note that onset features never count in rhyming. The diagram in (53) shows the speakers' intuition on this folksong.

(52)

| rhyme | aw | yaw | aw/yaw |
|---|---|---|---|
| speaker | A | C | G |
| | B | D | H |
| | | E | I |
| | | F | J |
| | | | K |
| | | | L |

Now let's look at a poem in which the syllables contain different pre-nucleus glides.

(53) tçyan li jye lyang **ÿÿan**
bu fu shi tshi **fan**
hong nyang çi tçyan çyan
ÿüe law lay tsheng **tçÿan**

'We have the chance to know each other, though we live far away.'
'It's worth it to know each other.'
'It's the matchmaker who makes us know each other.'
'It's Cupid who puts us together.'

For this poem, speakers had different intuitions from the first folksong. Among the same twelve speakers, six of them identified the rhyme as [an] (every line rhymes), two of them chose [ÿan] (only the first line and the last line rhymes), the other four recognized a gradient rhyme. Among the last four speakers, two feel that the first line and the last line rhyme closely, while the second line is okay as a rhyme. However, for the other two speakers, the first line and the last line are the best match, while the third line is also an acceptable rhyme. To account for the above four possibilities for rhyming, the only solution is to recognize the pre-nucleus glide as an independent constituent, which is part of the rime. For the first set of speakers, if there are different pre-nucleus glides present, rhyming only scans the R node. For the second set of speakers, rhyming always scans the R' node, so only syllables with exactly the same pre-nucleus glide can rhyme. The intuitions of the third set of speakers offers some support for the zero glide $\phi$. From the first folksong, we find that the rhyming constituent for speaker F is R'. In this poetry, we find that for this same speaker, the rhyming constituent still is the R' node, therefore, [ÿÿan] and [tçÿan] are the best match. The reason that [fan] is okay as a rhyme is because its surface representation is [f$\phi$an], while [çyan] is not okay as a rhyme is due to the fact that the pre-nucleus glide [y] is distinct from the other glide [ÿ]. From the first folksong, we find that the rhyming constituent for speaker L can be R or R'. This is consistent with his intuition in identifying the rhyme in this poetry. If the rhyming constituent is the R node, then the rhyme is [an] as predicted. If When the rhyming constituent is the R' node, the rhyme is identified as [ÿan] is also as predicted. The intuition of the last set of speakers, which overlaps to some extent with that of the third set of speakers, has implications for the current feature theory. The intuition of the last set of speakers suggests that labiality (or the feature [round]) is a secondary articulatory feature for vowels. For this set of speakers, two syllables may rhyme if they have the same pre-nucleus glide (in addition to having the same nucleus and coda), or if they have pre-nucleus glides with same specification for the feature [back]. Therefore, [ÿÿan] strictly rhymes with [tçÿan] and loosely rhymes with [çyan], but does not rhyme with [fan], [wϕan], or [dwan]. However, the latter three syllables do rhyme with each other. These results all support my proposal that pre-nucleus glides are in the rime. The following diagram shows the speakers' intuitions on this poetry.

| (54) | rhyme | an | ÿan | ÿan/an | ÿan/yan |
|---|---|---|---|---|---|
| speaker | | A | C | F | E |
| | | B | D | L | K |
| | | G | | | |
| | | H | | | |
| | | I | | | |
| | | J | | | |

To test native speakers' intutitions more accurately, I designed an experiment in which I presented each speaker with a song I constructed especially to probe properties of poetic rhyme. In this song, there are three occurrences of the pre-nucleus glide *y*, three occurrences of the pre-nucleus glide *ÿ*, three occurrences of the pre-nucleus glide *w*, three occurrences of words without any pre-nucleus glides, and also three occurrences of words that do not rhyme.[29]

(55) qing piao piao de giou shi guang giou zhe me liou zow
zhuan tou huei qü kan kan shi yi shu **nyan**
shi guang liou zhuang you duo sao bian **hwan**
fen bie hou si nian duei gi zai xin **kan**
tian ya hai giao ci qing he yi **gan**
ye ceng pan wang ÿe lao lai cheng **qÿan**
bu zhi suei ÿe neng fou ba si nian chong **dan**
hai shi yi ba ta bian zuo xi **gwan**

cang mang mang de tian ya lu shi ni de piao bo
gu dan dan de shen ying hou shi wo ji liao de xin qing
zuo ye meng li qing qing ba shou **wan**
wu liou liou de hei yan zhu he ni de xiao **lyan**
ni de shen ying bu duan di zai huei **xÿ̄an**
ru he ye nan wang ne rong **yan**
he shi cai neng ÿ̄e ÿ̄an ren **ÿ̄an**

'Time just flies.'
'It's been several years, now I look back.'
'Since then, how many changes there are.'
'I kept thinking of you after we separated from each other.'
'How can I stand this feeling that's with me wherever I am.'
'I have wished that Cupid could put us together.'
'But can time reduce my feelings toward you?'
'Or does time ingrain my feelings toward you?'
'Now you wander around the world.'
'I also feel alone and lonely.'
'Last night, I dreamed of holding your hands.'
'I saw your dark pretty eyes and your smiling face.'
'I kept seeing you in the dreams.'
'How can I ever forget your face?'
'When can we truly come together?'

Among ten of the same speakers as in the previous tests, one identified the rhyme as [an], and three identified four rhyming groups: [an], [ÿ̄an], [yan], and [wan]. Two recognized two rhyming groups: [an] and [wan] fall into one group, and [yan] and [ÿ̄an] fall into a second group. Four chose a gradient rhyme; they feel that it is better to identify four rhyming groups: [an], [ÿ̄an], [yan], and [wan], but it is acceptabl to choose [an] as the only rhyme. To account for the above four possibilities for rhyming, it is necessary to recognize the pre-nucleus as an independent constituent, which is adjoined to the rime.

A second kind of support for my hypothesis that the pre-nucleus glide is part of the rime, comes from Kejia, where the rhymes in the folksongs do not include the pre-nucleus glide (cf. R. Chung (1988)). However, Kejia differs from Mandarin in that two syllabl es can not rhyme if they contain different pre-nucleus glides. This gives support for my proposal that the pre-nucleus glide is part of the rime. If the pre-nucleus glide is in the onset, then one has to explain why syllables with different pre-nucleus glides can not rhyme.

The strongest counterargument to the idea that the pre-nucleus glides are part of the onset comes from Taiwanese. In Taiwanese, syllables with pre-nucleus glides seldom rhyme with those without the pre-nucleus glides, nor do they rhyme with those with different pre-nucleus glides (cf. Li (1986), Chang (1980)). If we recognize that pre-nucleus glides are part of the rime and recognize that there are two kinds of rhyming in Chinese, which we term "loose" and "strict" rhyming, then we can account for the different behaviors of the pre-nucleus glides in the different dialects. By proposing an independent constituent of R', we are able to set up the rhyming parameters for Chinese dialects. I argue that pre-nucleus glides are optional in rhyming in Mandarin, but obligatory in Taiwanese. Thus rhyming in Mandarin scans either the constituent R, or R', but rhyming in Taiwanese always scans the constituent R'.

## 6 Residual Problems

I have argued that pre-nucleus glides are part of the rime by drawing evidence from the distribution of mid vowels, distributional constraints on labial segments and back segments, language games, and rhyming. I find that Mandarin has a fixed syllabic structure of four slots, one in the onset and three in the rime. It remains to be shown whether all Chinese dialects share this same syllabic template.

A second question arises concerning the phonological status of the obligatory zero glide. The presence of this zero glide accounts for the Mandarin Labial Constraint without resorting to stipulatory device. It also accounts for the well-formedness of the lexical item /iai/ 'cliff' (with a surface representation of [yϕay]), which appears in Mandarin and Mandarin-based Fanqie languages, without resorting to exceptions to the Back Dissimilation Condition. In addition, it explains the behavior of both kinds of pre-nucleus glides in the Fanqie language of Mey-ka, without assuming that they have asymmetrical behavior, as Bao and Duanmu do. However, I do not see evidence of this sort in other dialects. Is the existence of the zero glide a special property of Mandarin and Mandarin-based Fanqie languages? I will leave these questions for future research.

## NOTES

*This paper is a revised version of a paper presented at the 25th International Conference on Sino-Tibetan Linguistics and Languages, 1992. I would like to thank Moria Yip, San Duanmu and audience at the conference for their useful comments.

[1] Note that the major difference between the proposal of Cheng (1973), Lin (1989), following the traditional syllable representation, and my proposal lies in that I claim that Mandarin has a fixed syllabic template of CGVX, instead of a canonical form, allowing one to maximally five segments

[2] Howie (1976) and Duanmu (1990) argue for the first process. Chao (1968) and Duanmu (1990) argue for (i) and (iv) of the second process. The arguments for the third process and for (ii) and (iii) of the second process are the contributions of this paper.

[3] However, in a weakened syllable, its duration is about 50% of that of a regular syllable. See Duanmu (1990) for more discussion of a weakened syllable.

[4] In my dialect, Mandarin spoken in Taiwan, only (6c and d) are possible variants of the zero onset for syllables with a [-high] nucleus. See Duanmu (1990) for detailed discussion of the zero onset.

[5] There is no data beginning with nonhigh vowels in Chao (1930), which should potentially give some insight to the 'zero onset' phenomenon.

[6] According to Li (1985), the zero onset in the Taiwanese Fanqie language has just one phonetic realization, i.e. the glottal stop [?]. This is true for every vowel-initial syllable. Taiwanese is unlike Mandarin, in that the Mandarin zero onset has four variants, and the Mandarin [+high] can spread onto the onset position.

[7] See section 3 for a more detailed description of a GV syllable in Mo-pa. Note that Chao (1930) does not include any data beginning with [w] for Mo-pa.

[8] Chao (1968, 18-23) divides the Mandarin syllable into four components, 'initial', 'medial', 'vowel' and 'ending', and suggests that initialless syllables have the 'zero initial', medialless syllables have the 'zero medial', and endingless syllables have the 'zero ending'. However, Chao does not provide explicit arguments to support his analysis.

[9] The non-occurrence of */fyan/ is probably just an accidental gap.

[10] Palatals are derived from velars, so palatals and velars are in complementary distribution.

[11] Note that /üai/ ([y̆ay]), and /üei/ ([y̆ey]) are not possible sequences either. Lin (1989) proposes two labial co-occurrence restrictions for Taiwanese, one operating on the syllable as a whole, and the other operating within the rime. Cheng (1989) and Duanmu (1990) criticize Lin's analysis, in that a syllable like *[kwaw] is ruled out twice. The same kind of criticism applies to my analysis, since a syllable like *[mwaw] is ruled out once by my proposed Labial Constraint, and again by the Back Dissimilation Condition. My point is that this type of ill-formed syllable is ambiguous, in that it is possible that both constraints are violated.

[12] This constraint also holds between the nucleus and the coda in Kejia in the case of a VG sequence. This constraint then precludes the following VG sequences in Kejia.

*ey *ow

In Cantonese, Kejia and Taiwanese, the mid vowels [e] and [o] appear in the underlying vowel inventory. Therefore, they are specified for the feature [back]. Thus, the fact that the combinations of [ye] and [wo] are bad in Kejia and Taiwanese is predicted by this constraint. However, it is not clear why [wo] is well-formed in Cantonese. Duanmu (1990) suggests that the two Labial nodes in this case merge into one, so that there is no violation of his proposed Round Constraint. We must then ask why merger does not operate in either Kejia, or Taiwanese. The two mid vowels [e] and [o] are derived from /6/ in Mandarin, and so are not specified for the feature [back]. They acquire their [back] and [round] specifications from [+high] neighboring segments, labial consonants, or from the default rule: [ ] → [+back]. Therefore, the combinations of [ye] and [wo] are good in Mandarin.

Note that coda consonants behave differently from post-nucleus glides in both Kejia and Taiwanese, where consonants are specified for the feature [back], and the nuclear vowel agrees with the coda consonant in [back] specification. For example, [ong], [ok], [ung], [ep], [ip], [im], and [em] are good combinations in Kejia, but *[om], *[op], *[um], *[up], *[ing], *[ik], *[eng], and *[ek] are all bad combinations. One could argue that the first four bad combinations are due to some type of Labial constraint, but it would be hard to account for the last four without resorting to some other device. See Chung (1988) for more detailed discussion. See also Chen (1990) for discussion of the different behavior of coda consonants and post-nucleus glides with respect to reduplication in secret language (Fanqie language) formation in Taiwanese.

Lin (1989) argues that the Final, which is equivalent to my proposed constituent R', is a domain for the Labial co-occurrence restriction and some other co-occurrence restrictions in Taiwanese. She finds that a high back vowel can not be followed by a velar consonant, therefore *[uk], and *[ung] are disallowed. The same kind of constraint also holds between the pre-nucleus glide and velar consonants in the coda, so that *uak and *uang are not allowed. She also finds that [ku] 'to squat' is well-formed in Taiwanese is because the two labial segments are not within the Final. However, she does not spell out what kind of constraint bans the combination of high back vowels and velar consonants. Here, I suggest that it is the same kind of Back Dissimilation Condition which operates within the R' as in Mandarin.

Nevertheless, it is not clear whether we can dispense with the Labial Constraint in favor of the Back Dissimilation Condition as the only constraint for Taiwanese, Kejia and Mandarin. The other alternative is to dispense with the Back Dissimilation Condition, and appeal to the Labial Constraint (Duanmu's Round Constraint) as the only constraint for these languages. This is exactly the approach Duanmu (1990) pursues. However, this proposal is not satisfactory either. For example, he must stipulate a restriction that primary labials (labial consonants) can not bear the feature [+round], in order to account for the badness of *[$b^w$a], *[$b^w$an], etc. in Mandarin. Similarly, he can account for the badness of *[uam], and *[uap] for Taiwanese, but only at the cost of stipulating [p] and [m] as secondary labials that bear [+round] specification in the coda position, so that the Round Constraint is violated. This stipulation is not otherwise motivated for Taiwanese. The same criticism can apply to Duanmu's analysis of Cantonese, in which he rules out *[wam], *[$k^w$am], *[wip], and *[$k^w$ip] by saying that [p] and [m] are secondary labials in coda position.

[13] /kya/ is realized as [tçya] as the result of palatalization. Besides palatalization, there are two other modification rules:
(24b') Insert a different tone onto the first rime (R), in case the first syllable of the output is totally identical to the base form (input).
(24c) Insert [l] into the second onset if the second onset of the input begins with [k].

[14] The rule $w \longrightarrow v/\#\_\_$ does not operate in all dialects.

[15] Another possibility would be to modify my proposed OSP so that any [+high] segment will be associated to the onset position first, then spread [+high] onto the pre-nucleus glide position, respecting the Labial Constraint. In that case, the Labial Constraint is formulated such that it can block the spreading of [+high] onto the pre-nucleus glide position.

[16] Note that May-ka and Mey-ka also differ in the application of palatalization. In Mey-ka, palatalization does not apply.

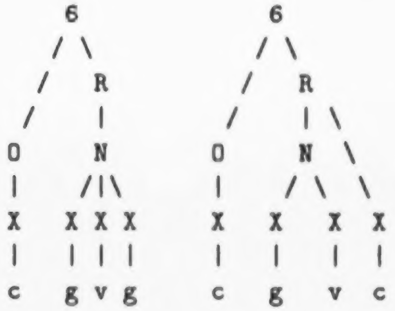
[17] Note again that Chao (1930) does not include any data beginning with [w] in GV syllables for Mey-ka. The advantage of my analysis is that I capture the generalization that the front glide [y] in syllable-initial position behaves the same in both May-ka and Mey-ka. As will be seen in (41) and (43), Bao's analysis provides no explanation for why the the same front glide has different behavior when preceded by the zero onset #.

[18] Data from the Taiwanese-based Fanqie language suggests that the coda consonant is outside the domain of application of substitution. Bao's proposed syllable structures given below in footnote 21, are consistent with the patterns observed in this Fanqie language. However, at this point it is still not clear whether substitution operates on the nucleus or on the rime constituent. What is certain is that substitution targets all the vocalic elements in a syllable, including both pre- and post-vocalic glides. It would require further study on other phonological processes to decide what the syllable structure is for this particular language.

[19] In this type of example, [n] shows up in the onset of the first syllable instead of [l], due to spreading of the nasal feature from the following vowel. [ĩ] shows up in the nucleus position of the second syllable because of nasality stability effect.

[20] There is no example beginning with [w], nor with a single vowel /i/ or /u/ as the only element in a syllable in Li (1985). However, there is another Taiwanese-based Fanqie language where the rules are: a. Reduplicate the syllable. b. Replace the onset of the second syllable with [l]. Here, I do find examples with /i/ or /u/ as the only element in the syllable besides the zero onset. For example, /i/ becomes as [?i-li] in this Fanqie language, while /u/ becomes [?u-lu].

[21] Bao (1990) proposes different syllable structures for different dialects. He claims that the pre-nucleus glide is part of an onset cluster in Mandarin, while it is part of the rime in dialects such as Mo-pa and Taiwanese. He proposes the the following two syllable structures for Taiwanese.



6 = syllable
O = onset
R = rime
N = nucleus

[22] Bao (1990) criticizes Yip's (1982) analysis, in that Yip derives correct output for [wan] and [yang] by associating [w] with a C slot, but associating [y] with a V slot. But nothing in Yip's theory can prevent associating [w] with the V slot, but associating [y] with the C slot. In the latter case, both of the ill-formed outputs *[way-kwan], and *[ye-kang] are derived. Note that Yip assumes that Mandarin and the Fanqie languages have a fixed skeleton CGVC, but allows some slots to be unfilled. She does not refer to syllable structure in her analysis.

[23] The following is Duanmu's formulation of Feature Recycling (Duanmu 1990:57):
a. Features and/or articulators in a floating/replaced segment may be reattached (i.e. recycled) to a nonfloating segment, without changing the existing features and articulators in the latter.
b. Which articulators/features to recycle is a language particular option.
c. Feature Recycling observes the phonotactics of the language in question.
d. Feature Recycling is local.

[24] Note that rule ordering is crucial in Duanmu's analysis. The w ⟶ v/#__ rule must apply before reduplication and substitution in order to derive the correct output. Otherwise, the incorrect output *[vay-k$^w$an] will be derived, since Feature Recycling will reattach the [+round] of the floating [w] back to the onset [k]. In contrast, there is no need to order the w → v rule before reduplication in either Bao's or my analyses.

[25] Duanmu (1990) does not recognize any back constraint for Mandarin. He believes that since [yay] 'cliff' is possible, the absence of CyVy is due to a gap. Duanmu (1992) (p.c.) suggests that he could also have a back constraint, encoded at the syllable node. In this case, however, his analysis would still suffer the same criticism as Bao's, sketched in section 4 where Bao's analysis is discussed. It seems that any analysis assuming that the pre-nucleus glide is part of the onset will have difficulties in accounting for glide-initial syllables in the language games.

[26] It's not straighforward to derive Palatalization in Duanmu's system.

[27] Duanmu (1990) suggests the following rules for Mo-pa and gives the derivation for [l$^y$ā] 'two' (Duanmu 1990:72).
a. Reduplicate the syllable.
b. Replace the first rime with [o]
c. Switch the value of [cont] of the second onset.
d. Simplify the first onset (i.e. delete the minor articulator).

| | |
|---|---|
| l$^y$ā | |
| l$^y$ā-l$^y$ā | (a) |
| l$^y$(ā)o-l$^y$ā | (b) |
| l$^y$(ā)o-t$^y$ā | (c) |
| l$^y$(ā)o-t$^y$ā | (Feature Recycling) |
| lo-t$^y$ā | (d) |
| output: lo-t$^y$ā | |

[28] Duanmu does not discuss the case of Mey-ka in his dissertation. However, we may construct an analysis along the lines of his analysis of May-ka. One possibility would be for him to say that there is no Onset Simplification operating in Mey-ka, but that the change from [l$^y$ey] to [ley] is due to Repair. Then he must account for why [yey] is good. This is exactly the same kind of dilemma he faces in his analysis of the Mandarin distribution facts, where he does not want to recognize two co-occurrence restrictions for Mandarin,

namely the Round Constraint and the back constraint. He chooses to say that the absence of *yVy is due to a gap. This raises the question of why the same kind of gap is also observed in other dialects, such as Kejia and Taiwanese. Is this really an accident? No such accidental, systematic gaps are required in my analysis. The other alternative for Duanmu's analysis would be to say that Feature Recycling only recycles [+round], not the feature [-back] in Mey-ka. This implicitly admits the asymmetrical behavior between the two kinds of glides. We mush then explain why the pre-nucleus glides in the two Fanqie languages, May-ka and Mey-ka, based on the same source language Mandarin, have such different behavior.

[29] The methodology of providing equal occurrences of the different pre-nucleus glides in testing native speakers' intuition was suggested to me by Chin-Chuan Cheng.

## REFERENCES

Bao, Z. (1990) Fanqie Languages and Reduplication, Linguistic Inquiry, 21.3 307-350.

Carreira, M. (1988) The Representation of Diphthong in Spanish, Studies in the Linguistic Sciences, 18, 1-24.

Chao, Y. R. (1931) Fanqie Yu Ba Zhong (Eight Varieties of Secret Languages Based on the Principles of Fanqie,) Bulletin of the Institute of History and Philology, 2.3, 320-354.

Chen, T. E. (1990) Implications of the Taiwanese Secret Language, Masters thesis, University of Illinois, Urbana-Champaign, Illinois.

Cheng, C. C. (1973) A Synchronic Phonology of Mandarin, The Hague, Mouton.

Cheng, R. L. (1966) Mandarin Phonological Structure, Journal of Linguistics, 2.2 135-158.

Chung, R. F. (1989) On the Representation of Kejia Diphthongs, Studies in the Linguistic Sciences, 19, 63-80.

Chung, R. F. (1989) Aspects of Kejia Phonology, Ph.D. dissertation, University of Illinois, Urbana-Champaign, Illinois.

Duanmu, S. (1990) A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages, Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Fudge, E. C. (1968) Syllables, Journal of Linguistics, 5, 253-286.

Li, F. K. (1966) The Zero Initial and the Zero Syllabic, Language, 42, 300-302.

Li, J. K. (1985) A Secret Language in Taiwan, Journal of Chinese Linguistics, 13, 91-121.

Li, J. K. (1986) Rhyming and Phonemic Contrast in Southern Min, Bulletin of the Institute of History and Philology, 57.3, 439-463.

Lin, Y. H. (1989) Autosegmental Treatment of Segmental Processes in Chinese Phonology, Ph.D. dissertation, The University of Texas at Austin.

Liu, N. H. (1944) Chengdu Erton Jian de Mimi Yu, (The Secret Language of the Children in Chengdu,) Bulletin of Chinese Studies, 4.2, 69-78.

Marantz, A. (1982) RE Reduplication, Linguistic Inquiry, 13, 483-545.

Norman, J. (1988) Chinese, Cambridge University Press, Cambridge, UK.

Steriade, D. (1988) Reduplication and Syllable Transfer in Sanskrit and Elsewhere, Phonology, 5.2, 73-155.

Wang, L. (1973) Hanyu Shilu Xue, (The Chinese Versification,) Zhonghua Shuju, Shanghai.

Yip, M. (1982) Reduplication and C-V Skeleta in Chinese Secret Languages, Linguistic Inquiry, 13, 637-661.

Yip, M. (1989) Cantonese Morpheme Structure and Linear Ordering, paper presented at WCCFL, University of British Columbia, Vancouver, Canada.

# VENDLER CLASSES AND REINTERPRETATION[*]

Michihiko Kawamura
University of Texas at Austin

Abstract: This paper is concerned with the theoretical status of Vendler classes in grammar and utterance interpretation. I argue that Vendler classes should be reinterpreted as a classification of semantic objects in our aspectual ontology and that we need only four discrete categories. Based on this categorization, I propose a relevance-theoretic account of aspectual interpretation in context. The framework sketched in this paper presents a simple representation of English aspectual system and leads to the conclusion that the apparent complexity and subtleties of the aspectual phenomena are epiphenomena created by the interaction between grammar and cognition.

## Introduction

Verbal aspect has long been a topic of interest in both linguistic and philosophical traditions. Among the numerous studies in this area, Vendler (1957) is particularly important to us, because most of the recent works in linguistics are carried out making reference to his classification system. This paper aims to clarify the nature of aspectual information encoded in the lexicon of English based on the assumption that aspect is a purely temporal phenomenon. The first half of this paper discusses Vendler's (1957) classification of lexical aspect. Though most recent works on this topic simply assume his four-way taxonomy, it has been frequently pointed out that his classification has some difficulties. Many refined versions classify verb phrases or atomic sentences into more than four categories. These refined classifications are, however, by no means unchangable. In many cases, a verb can shift from its most 'natural' category to another quite easily, and seemingly ungrammatical sentences become acceptable with a little imagination. Considering the vast amount of work on aspect, relatively few attempts have been made to account for this 'reinterpretation' problem (Lys & Mommer 1986, Moens & Steedman 1988, Smith 1992). I present an alternative theory that makes a clear distinction between semantic objects and their names on one hand, and between semantic meaning and pragmatic meaning on the other, and I argue that Vendler's taxonomy is linguistically significant as classification of situations denoted by linguistic expressions, but not of expressions themselves. In this analysis, only four eventualities on the line of Vendler's are to be distinguished to represent the temporal constitution of situations

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 53-88

denoted by predicative expressions, and the boundary between these four categories are clear-cut.[1]

The latter half of this paper handles the 'reinterpretation' phenomena by drawing a clear boundary between grammar and cognition, stating how aspectual semantics fits into the theory of utterance interpretation. A closer examination of the 'reinterpretation' phenomena reveals that the different categorizations of eventuality instruct us to understand linguistic expressions differently via such processes as metonymical extension. Together with the assumption that lexical aspect is purely and simply a notion of time, the framework sketched in this paper presents a simple representation of the English aspectual system and leads to the conclusion that aspectual semantics is not so complicated as generally considered.

Before leaving this introductory section, I would like to make clear the semantic framework on which this informal study will be based. The basic framework employed here is model-theoretical semantics, semantics as a correspondence theory between language and something that is not language. More specifically I will adopt Bach's (1986b) view of semantics as natural language metaphysics. Following Bach, I will distinguish the theory of meaning as part of the grammar from the theory of utterance interpretation as part of our cognition. I also assume aspect as a purely temporal phenomena, which is shared and motivated by algebraic semantics proposed in Link (1983, 1987), Hinrichs (1986) and Krifka (1986, 1989, 1992a). As for the pragmatic/cognitive framework, I will employ the Relevance Theory developed by Sperber and Wilson (1986). Details of this theory will be introduced when necessary in the discussion.

## 1. Vendler (1957) and Beyond

### Vendler's (1957) Classification

Most of the recent works on lexical aspect are carried out either as a reaction to or by making reference to Vendler's (1957) classification. He set out a taxonomy of verbs, where VERBS are classified into FOUR aspectual classes of EQUAL status. The four categories distinguished are given in (1) with representative examples of verbs in each category.

(1) a. activity terms: *run, walk, swim, push (a cart), drive (a car),* etc.
    b. accomplishment terms: *paint (a picture), make (a chair), build (a house), run (a mile), walk (to school), deliver (a sermon),* etc.
    c. achievement terms: *reach (the summit), win (the race), die, find,* etc.
    d. state terms: *have, desire, love, hate, want, know, believe, rule,* etc.

I follow his terminology throughout this paper. Vendler employs several linguistic tests to distinguish these verb classes.[2] First, based on their ability to occur in the progressive, Vendler distinguishes activities and accomplishments from states and achievements. Assuming that the progressive form is compatible only with durative situations, he characterizes the former as 'continuous' and the latter as 'punctual'.

(2) What are you doing
    a. I am running. (activity)
    b. I am writing a letter. (accomplishment)
    c. *I am recognizing Jim. (achievement)
    d. *I am knowing Mary. (state)

Next, activities and accomplishments are distinguished by the kind of adverbials they are compatible with. Durational adverbials with *for* occur with activities, but not with accomplishments (Test 2). Just the opposite is the case with adverbials with *in* and *take* ... *to* ... construction (Test 3). Based on these observations, Vendler designates activities and accomplishments as temporally homogeneous and heterogeneous respectively. In other words, accomplishments do and activities do not have a set terminal point 'which is logically necessary to their being what they are' (Vendler 1967: 101).

(3) a. For how long did he push the cart? (activity)
    He pushed the cart for half an hour.
    b. *How long did it take to push the cart?
    *It took for half an hour to push the cart.
    *He pushed the cart in half an hour.

(4) a. *For how long did he draw the circle? (accomplishment)
    *He drew the circle for twenty seconds.
    b. How long did it take to draw the circle?
    It took twenty seconds to draw the circle.
    He drew the circle in twenty seconds.

States and achievements are also distinguished from the kinds of compatible temporal adverbials. States are compatible with *for*-durationals, but not with punctual adverbs. On the other hand, achievements occur with instantaneous adverbials only. Vendler notes that 'while achievements involve unique and definite time instants, states involve time instants in an indefinite and nonunique sense.' (Vendler 1967: 107). His classification system is summarized in Table 1 below.

(5) a. At what time did you reach the top. (achievement)
    I reached the top at noon sharp.
    b. *For how long did you reach the top?
    *I reached the top for three years.

(6) a. *At what time did you love her? (state)
    *I loved her at noon sharp.

b For how long did you love her?
I loved her for three years.

| verb class \ test | Test 1 | Test 2 | Test 3 | Test 4 |
|---|---|---|---|---|
| Activity | OK | OK | * | n.a. |
| Accomplishment | OK | * | OK | n.a. |
| Achievement | * | * | n.a. | OK |
| State | * | OK | n.a. | * |

Test 1: Ability to appear in the progressive
Test 2: Ability to occur with temporal adverbials headed by *for*
Test 3: Ability to occur with temporal adverbials headed by *in*
Test 4: Ability to occur with instantaneous adverbials
(n.a. = the test is not applied by Vendler)

Table 1: Vendler's (1957) classification tests

**Post-Vendler Classifications**

In Vendler's analysis, VERBS are classified into FOUR categories of EQUAL standing. It has been frequently pointed out, however, that his classifications have some problems, and many refinements have been proposed. This subsection summarizes some of the main problems with Vendler's classifications and new directions proposed in the literature.

Structured taxonomy: Vendler's aspectual categories must be structured taxonomically. In other words, the four categories are not of the equal standing with each other. Mourelatos (1978) and many others argue that Vendler fails to make the basic distinction between states and non-states on the one hand, and between telic and atelic events on the other. Contra Vendler, achievements are more like accomplishments than states, as illustrated by the following example.

(7) The patient was dying in a few minutes. (achievement)

Based on such observations, Mourelatos (1978) introduces a structured taxonomy, which is reproduced here with my terminology in part:

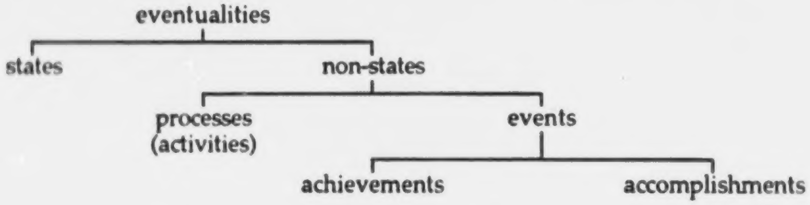


Figure 1: Mourelatos' (1978) taxonomy

Inflation of categories: It is frequently pointed out that Vendler's four-way taxonomy is not exhaustive, and many refined versions of

Vendlerian classification include more categories than Vendler's original. Dowty (1979), Moens and Steedman (1988) and Smith (1991) distinguishes five categories, L. Carlson (1981) six, Bach (1986) seven, Lys and Mommer (1986) seven. For example, Bach (1986) proposes the following classification, reproduced here with our terminology in part.

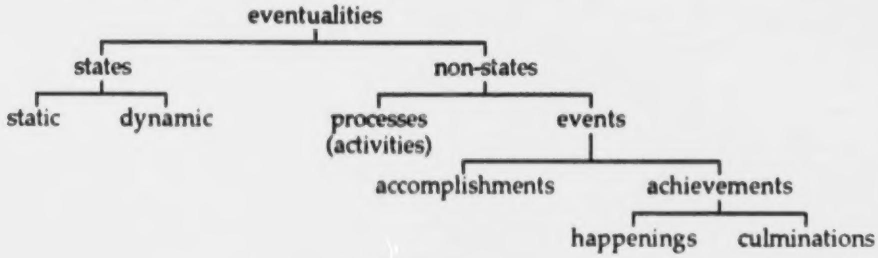


Figure 2: Bach's (1986) taxonomy

The question remains as to how many classes are to be distinguished. We will return to this problem later in this section.

Aspectual composition: Verbs are not the only determinant of aspectual meaning, as pointed out in Garey (1957), Verkuyl (1972), Mourelatos (1978), Platzach (1979), Dowty (1979), L. Carlson (1981) and elsewhere.

(8) a. John ate an apple yesterday (accomplishment)
John ate apples yesterday. (activity)
b. John ran this morning. (activity)
John ran to the station this morning. (accomplishment)

For example, sentences with the transitive verb *eat* can be an accomplishment or an activity sentence depending on the type of its object NP. Similarly, the adverbial *to the station* takes an activity verb *run* and return an accomplishment VP *run to the station*. It is now clear that the aspectual types are not predictable from the verb alone. However important, the verb is only a determinant of sentential aspect, and therefore it has become a widely accepted idea that it is not the aspectual properties of verbs but those of VPs or atomic sentences that must be classified.[3] It does not seem linguistically significant, however, to classify sentences in terms of their aspectual meanings. Rather, we are in need of a theory of aspectual composition that can predict the sentential aspect from the meaning of its constituents. In fact, some attempts have been made, for example, in Verkuyl (1972), Platzach (1979), Mittwoch (1980), Krifka (1989, 1992), Tenny (1987), and Smith (1991). In this paper, we are not concerned with the nature of aspectual composition in any detail.

In aspectual composition, we need to make reference to lexical information of verbs. In this context, however, Vendlerian verb classification does not help so much. We cannot predict aspectual properties of complex expressions by simply marking a verb with one of Vendler's four labels. Introducing more eventuality labels would not make this situation any better, but would merely make aspectual semantics more complicated. In short, Vendlerian verb classification fails when it identifies the set of aspectually distinguishable verb classes with the set of eventuality types. Possible aspectual meaning of a verb is predictable from its non-aspectual meaning, and verbs of a semantic class usually share the same set of possible aspectual meaning. This is not visible in Vendlerian classification. Furthermore, many verbs are polysemous, and verbs of different semantic classes show different syntactic alternation patterns (See Miller & Fellbaum 1991, Levin 1993 for detailed discussion). Since verbs' thematic structure and their syntactic realization pattern affect their aspectual meaning, verbs under the same Vendlerian class may have different sets of possible aspectual meanings and their composition because of their difference in meaning. Viewed in this light, we safely conclude that Vendlerian verb classification is too simplistic to be useful in the aspectual composition. What we need is not a simple tree-like structure but a more refined classification based on independent semantic motivations.

Reinterpretation phenomena: Thus far, we have looked at some directions in the study after Vendler (1957). The most distant version from Vendler will be the one that classify linguistic expressions hierarchically into more than four categories. The refined versions of aspectual classifications are, however, still not secure enough. In many cases, a verb or a sentence can 'shift' from its most 'natural' category to another quite easily. Seemingly ungrammatical sentences start making sense with little imagination, which is one source of the common assumption that aspectual semantics is complicated and elusive. Following sentences illustrate this point.

(9) a. John ran at three o'clock. (activity → achievement)
b. At 3 o'clock, Mary knew the answer. (state → achievement)

Given a context, these sentences become perfectly natural. When we want to make our aspectual semantics plausible, we need to tackle this reinterpretation phenomenon, instead of using it as a wastebasket. I will tackle this problem later in Section 3.

## 2. A Neo-Vendlerian Approach

### Introduction

It was observed in the previous section that Vendler's classification should not be regarded as linguistically significant as a classification of verbs and other linguistic expressions. Rather, we need a more fine-grained verb classification motivated by more general grammatical considerations. Is Vendler's view too simpleminded to provide the basis for aspectual semantics? My answer is 'No!'. In this section, I present an alternative theory that, I think, comes closer in spirit to Vendler's original.

First, I reinterpret Vendler's classes as types of aspectual situations denoted by linguistic expressions, but not classes of linguistic expressions themselves:[4]

(10) CLASSIFICATION OF SEMANTIC OBJECTS:
Vendler classes are categories of semantic objects denoted by linguistic expressions.

Then, I go on to argue that only four aspectual classes on the line of Vendler are to be distinguished to represent the internal time of linguistically described situations, and the category boundaries are not fuzzy:

(11) EXHAUSTIVENESS OF VENDLER CLASSES:
All and only the four categories proposed by Vendler are relevant for the aspectual classification of eventualities. We understand linguistically described situations as instantiating either one of Vendler's four situation types but nothing else.

I modify Vendler's taxonomy slightly by making the event-state distinction as the primary one.

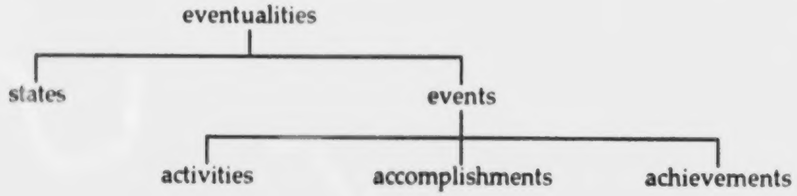


Figure 3: Classification of Eventualities

In this framework, the apparent complexity and subtlety in aspectual semantics is the epiphenomena created by the interaction of aspectual semantics with other aspects of grammar and human cognition. The 'reinterpretation' phenomena is handled by drawing a clear boundary between grammar and utterance interpretation, stating how extra factors enter into aspectual semantics. Together with the assumption that lexical aspect is purely and simply a notion of time, the framework sketched in this paper presents a simple representation of the English aspectual system.

To show how this system works, I take two steps. First, after reviewing the temporal nature of the four Vendler classes, I will illustrate that we need no additional categories. Our eventuality ontology will be made clear and simple by putting all the leftovers into pragmatics and other linguistic wastebaskets. Then, we turn to the 'reinterpretation' phenomena and other semantics-pragmatics clashes, cleaning up the pragmatic basket.

**Aspectual Classification of Semantic Objects**

Let us turn to look at the temporal nature of Vendler's four situation types, namely states, activities, accomplishments and achievements. These four types are distinguished from one another by two tests which are purely linguistic and temporal in nature: compatibility with point-time adverbials and durative adverbials.[5] Note that asterisks indicate only that the sentence is ungrammatical in the intended reading, which is semantically ill-formed in the framework we are assuming. Other possible readings are discussed in Section 3.

Simple eventualities: Let us start with the characterization of events. Events occur at some particular time interval. For achievements the interval is a point in time, while for activities and accomplishments it is a time period.[6] In other words, achievements are punctual and activities and accomplishments are durative. The difference in temporal length of events is reflected in the compatibility with temporal adverbials.[7] Thus, achievements can occur with point-time adverbials while durative events cannot.[8]

(12) a. *The kids played at 3:00. (activity)
b. *John ran a mile at 3:00. (accomplishment)
c. Fred crossed the border at 3:00. (achievement)

This indicates that activities and accomplishments expressions both denote durative events. However, activities and accomplishments differ in their internal constitution. Accomplishments are inherently bounded (telic) while activities are not (atelic). In other words, accomplishments have a distinct culminating point as well as a durative process, but activities are homogeneous processes without culmination.[9] In this sense, accomplishments consist of an activity followed by an achievement.[10] This difference in temporal constitution determines the compatibility with durational adverbials with *for*:[11]

(13) a. The kids played for three hours. (activity)
b. *John ran a mile for three hours. (accomplishment)
c. *Fred crossed the border for three hours. (achievement)

In sum, activities are durative events that consist of homogeneous subevents. Accomplishments are heterogeneous events consisting of homogeneous subevents and a distinct culmination point. Achievements are atomic events which are unanalyzable into subevents.

| | +homogeneous | –homogeneous |
|---|---|---|
| +durative | activity | accomplishment |
| –durative | achievement | — |

Table 2: Classification of Events

Having an built-in endpoint, heterogeneous events have what Manfred Krifka names 'quantized reference property': whenever there is an event to which an accomplishment predicate applies, then no proper part of it can be an event described by the same predicate. Formally,

$$\forall P\,[\mathrm{QUA}(P) \leftrightarrow \forall x, y\,[P(x) \wedge P(y) \rightarrow \neg\, y \sqsubset x]] \qquad \text{(Krifka 1992a: 32)}$$

For example, no proper part of an event of running a mile can be an event of running a mile. Accomplishments have quantized reference as well, since they are atomic events without internal constitution. Activities have, on the other hand, cumulative reference properties: whenever there are two different events named by an activity predicate, then the same predicate can be applicable to the sum event (cf. Quine 1960). See Krifka (1992a) for lattice-theoretic characterization of activities and accomplishments.

Next, we turn to distinguish states from events. As pointed out elsewhere, this is the primary distinction.[12] States hold, instead of occurring at some particular intervals. They sometimes appear to pattern with some event types, but they always behave differently. First, states can occur with point-time adverbials like achievements but differ in their interpretation.

(14) a. At 12:00, John was angry. (state)
b. At 12:00, John arrived. (achievement)

In (14a), the atomic sentence *John be angry* holds true for an indefinite duration including the instant specified by the adverbial. In (14b), on the other hand, does not allow this interpretation. While event sentences can be true only at some particular intervals, state sentences allow for a durative interpretation in which the situation holds for an indefinite duration (cf. Vlach 1981).[13] In this paper, we are not concerned with the details on the nature of states and stative expressions. I simply assume that events are bounded by temporal endpoints while states are not.

| –bounded | +bounded |
|---|---|
| states | events |

Table 3: Events and States

The following example, due to Barbara Partee, appears to be a counterexample to this characterization (C. S. Smith p.c.). In (15), the temperature was ninety only briefly, perhaps only at noon.

(15) At noon, the temperature was ninety and rising.

The durative interpretation is, however, necessary here. Suppose that we observed that the temperature was ninety at 12:00 sharp, and it rose to ninety-one after a second when we observed the thermometer again. It is still possible for the state expression *the temperature be ninety* to be true at 30 milliseconds after 12:00, since states are unbounded in time. Events are not, however, interpretable this way. When an event expression (e.g. *John arrive*) is true with respect to some interval (e.g. 12:00), the particular event cannot hold true at the same time with respect to another interval (e.g. thirty seconds after 12:00).

Second, states are like activities in that they occur with *for*-durationals. There seems to be no qualitative difference between the two situation types when bounded by durational adverbials, but this is not true. When bounded by *for*-durationals, activities merge into accomplishment, while states do not (cf. Bach 1981). The difference becomes visible, for example, when they appear in the *It take ... to ...* construction.

(16) a. Naturally, it took John an hour to run for an hour.
b. *Naturally, it took John a year to love Mary for a year.

Complex eventualities: I distinguish four types of complex eventualities, namely ITR(ϕ), PROG(ϕ), PERF(ϕ) and GEN(ϕ), which correspond to iterative, progressive, perfect and generic/habitual sentences respectively.

First, event sentences denote activity-type events under an iterative interpretation.

(17) a. John sang two songs in a minute. (accomplishment)
b. John sang two songs for three hours. (activity)

(18) a. Mary kicked the ball at that moment. (achievement)
b. Mary kicked the ball for an hour. (activity)

Iteratives are a subcategory of activities. They contain 'atoms', and are like nominal plurals in this respect (e.g. Krifka 1989, 1992a). To represent this, I employ a phonologically null predicate ITR that takes an event to form a complex predicate of activity type.

(19) SEMANTICS OF ITERATIVE PREDICATE
a. ITR maps an event into its iterative counterpart
b. ITR($\phi$) denotes an activity

Though the iterative reading is less salient in activities and accomplishments than in achievements, I consider all events to be inherently ambiguous at least in the temporal semantics of aspect. That is, the ITR operator applies blindly to all event types.

First, iteratives need to be separated from quantized events. Quantized events are modified by 'cardinal count adverbials' (i.e. *once, twice, N-times*) and are count-quantified in their 'nominalization transcriptions' (Mourelatos 1978). Iteratives are, on the other hand, mass-quantified in their nominal counterpart.

(20) a. John kissed Mary twice yesterday. (quantized)
b. There were two kissings of Mary by John yesterday.

(21) a. John ran a lot yesterday. (iterative)
b. There was a lot of running by John yesterday.

The paradigm in (20)-(21) suggests that quantized events are heterogeneous, needing to be distinguished from homogeneous iteratives. In fact, quantized events show the expected behavior with respect to atelic *for*-durational adverbials:

(22) a. [John kicked the ball $_{\text{ATELIC}}$] for three hours. (iterative)
b. *[John kicked the ball twice $_{\text{TELIC}}$] for three hours. (quantized)

Next, iterative readings need not be regarded as special, in spite of the fact that the interpretation is not always available in accomplishments and achievements. (The # marks are used to indicate that the sentence is unacceptable or ungrammatical for some reason we are not concerned within the current framework.)

(23) a. #Jack built a house for three days. (iterative)
b. #John killed an old man for three hours.

In (23), the iterative reading is impossible. However, it is not the aspectual semantics that makes these sentences anomalous. In fact, iteratives are impeccable with bare plurals and intensional NPs, as shown in (25a) and (b) respectively. In (25b), there must be two different houses that Jack built. Iteratives are also possible in fiction and other possible worlds where our normal assumptions about the real world do not hold, as (25c) and (d) illustrate. The contrast between (24) and (25) suggests that we should not to mark sentences like (23) and (24) as ungrammatical in aspectual semantics. Rather, they must be filtered out by some other mechanism in our lexical knowledge of language and the world.

(24) a. #Jack built a house twice.
b. #John killed an old man twice.

(25) a. During his stay in Tokyo, John ate hamburgers three times.
b. Jack built David's house twice.
c. John killed the zombie three times.
d. In his dream, John killed an old man twice.

If the line of argument presented here is on the right track and accomplishments and achievements are inherently ambiguous between singular and iterative event readings, they must allow mass quantification in nominalization transcriptions. This prediction is borne out, as illustrated by the following example due to Bach (1989).[14]

(26) There $\left\{\begin{array}{l}\text{a. were lots of horses}\\ \text{b. *was a lot of horse}\end{array}\right\}$ in the stable. (count)
There $\left\{\begin{array}{l}\text{c. were lots of kissings}\\ \text{d. was a lot of kissing}\end{array}\right\}$ during the parade. (telic/atelic)

We can therefore conclude that all event verbs are potentially ambiguous between simple and iterative readings and the application of ITR is context-free to all event types. One advantage of this analysis of iteratives is that we do not have to state a category-shift rule that makes reference to the subject NP in order to derive the iterative activity reading. With the assumption that one and the same bomb does not explode twice, (27a) has only one single-event reading, but in (27b) the iterative reading is possible with the distributive reading of the subject NP *the bombs*. Thus, (27c) is fine when only a part of the bombs exploded.

(27) a. The bomb exploded. (achievements)
b. The bombs exploded. (activity/achievement)
c. The bombs exploded for an hour.

I also suggest three operators that form complex states. The PROG and PERF operators map eventualities into stage-level states, and the GEN operator maps events into individual-level states.[15] The details about these operators are beyond the scope of this paper.

(28) a. John wrote a letter. (event)
b. John was writing a letter. (progressive: stage-level state)
c. John has written a letter. (perfect: stage-level state)
d. John writes a letter every day. (generic: individual-level state)

### *Back to Vendler: Reducing Aspectual Categories*

In the classification systems proposed hitherto, predicates are classified into several classes directly related to eventuality types, with possible later category conversions by rules. To put it different way, even when classifications are made of situations, they are at the same time classifications of linguistic expressions associated with the given eventuality type in one-to-one fashion.[16] For example, Lys and Mommer (1985) (henceforth L&M) present an extreme example of aspectual classification of this kind. L&M distinguish seven types, and Inagaki and Iwata (1990) show that seven is not enough and expand their categories into eight, still not being able to distinguish accomplishments properly. My claim in this paper is that the classification of eventualities should be independent of classification of linguistic expressions, and we need all and only Vendler's four categories. The classification of linguistic expressions like VPs or Ss does not seem to be linguistically significant but just encyclopedic. Rather, classifications should be made of lexical items in terms of their inherent (possibly underspecified) aspectual value and the possible compositional processes they may undergo. However, this paper is not concerned with this problem of how the lexicon and the grammar interact with each other to compose the sentential aspect. See Tenny (1986), Verkuyl (1989), Smith (1991) for detailed discussion of aspectual composition. The remainder of this section is devoted to the illustration of how extra categories other than Vendler's can be eliminated.

Dynamic states: Let us start with 'dynamic' verbs. In many aspectual classifications, such verbs as *stand* and *lie* (henceforth dynamics) are given an independent status (e.g. Dowty 1979; L. Carlson 1981; Quirk et al. 1985; Bach 1986). It is claimed that these verbs are between states and activities from the aspectual point of view. L. Carlson (1981) points out that they are like states because they occur with instantaneous adverbials, but like activities in that they can appear in the progressive.

(29) a. At seven o'clock, the caravan stood in its old place. (state)
b. At seven o'clock, the caravan was standing in its old place. (activity)

Notice that the subject in (29), *the caravan*, is ambiguous between animate and inanimate, or possibly between agentive and non-agentive. Ambiguities of this sort can hold in Japanese as well. In Japanese, nouns like *kyaraban* 'caravan' can occur not only in *aru*-existentials that take inanimate subjects but also in *iru*-existentials that take animate subjects, as shown in (30) below:

(30) a. Kyaraban-ga motono-basho-ni atta
caravan-NOM old-place-LOC be$_{aru}$-PAST
'The caravan stood in its old place'

b. Kyaraban-ga motono-basho-ni ita
caravan-NOM old-place-LOC be$_{iru}$-PAST
'The caravan was standing in its old place'

These examples suggest that it is possible to argue that verbs like *stand* and *lie* appear either as an activity when they occur with an animate subject or as a state when they occur with an inanimate subject.[17] In fact, as I will argue, this polysemy analysis is preferable. Consider the following. According to Quirk et al. (1985), dynamics are 'characterized by their ability to be used both (a) with the nonprogressive to express a permanent state, and (b) with the progressive to express a temporary state'.

(31) a. The city lies on the coast.
His statue stands in the city square.
b. People are lying on the beach.
He is standing over there.

However, the contrast Quirk et al. (1985) found is only superficial. First, the permanent-temporal contrast is not unique to dynamic verbs under our assumption of the function of the progressive. Application of the progressive to a state sentence yields a temporal interpretation.

(32) a. Mary eats Japanese food. (state: generic)
b. Mary is eating Japanese food these days. (state: progressive)

(33) a. John lives in Austin. (state: lexical)
b. John is living in Austin. (state: progressive)

Second, the permanent-temporary distinction in (31) corresponds to the subject's inanimate-animate distinction. If dynamic verbs denote a single situation type, then they are expected to behave with no regard to the subject's animacy. Contrary to this expectation, these verbs behave differently when making reference to a subject's animacy:

(34) dynamics referring to an activity
a. *At that point John stood in front of the Tower.
*Mary lay on the ground at that moment.
b. John stood in front of the Tower for three hours.
Mary lay on the ground for a while.
c. John was standing in front of the Tower for three hours.
Mary was lying on the ground when John arrived.

(35) dynamics referring to a state
a. The statues stand in front of the Tower.
b. The statues are standing in front of the Tower.
c. *The statues stand in front of the Tower at this very moment.

The paradigm in (34) indicates that *stand* and *lie* denote activities with human subjects. In contrast, (35) illustrates that these verbs name states with inanimate subjects. In other words, dynamic verbs are polysemous lexical items. The point is: dynamics cannot denote both an activity and a state at the same time, but rather they have interrelated multiple senses with different aspectual meanings.

(36) a. John ran through the town.
The river runs through the town.
b. Mary was facing an invader.
Her house faces the river.
c. John looked at the picture.
His room looks to the south.

Note that dynamic verbs do not form the only verb group that denotes both events and states. Many other verb classes including perception and psych verbs have this multiple-entry property as well. Dynamic verbs are not different from them in this respect.

Result states: 'Result state' verbs can take *for*-durationals that modify result state of the event named by the verb (Lys and Mommer 1985).

(37) a. The janitor unlocked the door for fifteen minutes.
b. The lake froze for forty days.

What should be noted here is that *for*-durationals in (37) specify durations of the result state which by itself is not named by the verb, and have no influence on the determination of eventuality named by the verb.[18] These result states, as Mittwoch (1980) points out, do not necessarily hold after the event expressed by the sentence. This means (38) is not an accurate paraphrase of (37a). This becomes visible when the subject is agentive. Thus the sentences in (39) contains nothing contradictory:

(38) The janitor unlocked the door and the door was unlocked for fifteen minutes.

(39) a. He lent me the book until Monday but I gave it back on Sunday. (Mittwoch 1980)
b. Yesterday John went to Tokyo for a week.

These examples indicate that the result states are intensional and therefore not located on the real-time axis together with their causal events as the states actually holding in the world. For this reason, I take result states simply as states that are introduced as an implicature in this special construction. I assume this construction as an instance of the English resultative construction. See Goldberg (1992) for detailed examination of the resultative and related constructions in English.

**Heterogeneity-neutral events**: 'Unspecified culmination point' verbs can occur with both telic and atelic adverbials (Lys & Mommer 1985).

(40) a. The icecream corn [sic] melted in/for ten minutes.
b. Grandmother Lina washed the skillet in/for five minutes.

They can be either accomplishments or activities, depending on the speaker-hearer's assumption, i.e. according to whether s/he views the change as relative or absolute. However, they cannot denote both at the same time.[19] These verbs are underspecified for heterogeneity. Inchoative achievements may therefore form a verb class, but may not form a verb class that is uniquely related to some eventuality.

**Duration-neutral events**: Lys & Mommer (1985) analyzed the verbs in (41)-(42) as denoting DN-CP event (durative nucleus + culminating point). They can take both point-time adverbials and telic adverbials.

(41) The janitor locked the door $\left\{\begin{array}{l}\text{a. at 10 o'clock}\\ \text{b. in ten minutes}\end{array}\right\}$.

(42) The cat drowned $\left\{\begin{array}{l}\text{a. at 10 o'clock}\\ \text{b. in ten minutes}\end{array}\right\}$.

The temporal scheme of DN-CP appears to correspond to accomplishments, but there is one important difference which L&M fails to distinguish:

(43) *John built a house at 3:00. (accomplishment)

If the temporal organization of these is the same, they should behave in the same manner aspectually. Otherwise, we need to posit another situation type to make them different, which will complicate our aspectual semantics. Furthermore, a closer examination of (41)-(42) will tell us that these verbs do not even form a natural class.

(44) a. *The janitor locked the door in ten minutes. It was at 10:00.
b. The cat drowned in ten minutes. It was at 10:00.

The event of locking the door can be durative or instantaneous, depending on the context, while the event of drowning is not. The verbs like *lock* must be underspecified for duration. They can be either accomplishment or achievement but never both at the same time. On the other hand, verbs like *drown* cannot be durative. This is true whichever possible world we are in, whatever belief we have. Verbs like *drown* should not be treated as accomplishments because (42b) does not mean that the cat's drowning continued for 10 minutes.

For this reason, instantaneous change-of-state verbs like *drown, arrive, die* and *reach (the top)* (henceforce, *arrive* verbs) are frequently isolated from other verbs which denote instantaneous events (e.g. L. Carlson 1981; Bach 1986; Moens & Steedman 1987, Talmy 1988; Smith 1991). It is argued that they denote a different type of eventualities from verbs like *knock, blink,* and *hit* (henceforce, *knock* verbs). Consider the contrast between (45) and (46). While *knock* verbs in (45) are often interpreted as denoting iterative events, *arrive* verbs in (46) normally appear in the progressive as denoting a single event. This difference in interpretation is another motivation for the special treatment of *arrive* verbs.

(45) *arrive* verbs: Prog($\phi$): $\phi$ = accomplishment
a. John is arriving in five minutes.
b. The patient was dying.
c. Mary was reaching the top.

(46) *knock* verbs: Prog(ITR($\phi$)): $\phi$ = achievement
a. Jenny was knocking at the door.
b. Bill is blinking now.
c. Mary is hitting John.

Suppose that the eventuality denoted by *arrive* verbs has an independent status. Then, we need to explain why they can take both instantaneous and durative adverbials as in (42). I argue that the verb in (42) simply names an achievement and, does not instantiate a special event type. But, then, a question arises as to why (41d) is possible. Remember that we are separating the classification of eventualities as types of semantic objects from the classification of linguistic expressions as names of these semantic objects. My answer is that it is possible via 'reinterpretation'. Similarly, we can simply state that the sentences in (45) are the progressive of (reinterpreted) accomplishments. We will discuss such 'reinterpretation' procedures in Section 3.

In this section, I have argued that if we aim at a classification of eventualities in terms of a purely temporal point of view, we can avoid inflation of aspectual categories and limit them into Vendler's original four. This means that semantic objects denoted by linguistic expressions fall into one of the four categories without any exception, and leads to a rather strong hypothesis about human cognitive activity that people categorize situations in the world in one of the four types.

## 3. Vendler Classes in Contexts

### Grammar vs. Pragmatics

Thus far, we have examined Vendler's four-part taxonomy as temporal classification of semantic objects denoted by predicative expressions from a purely temporal viewpoint ignoring extratemporal as well as extralinguistic factors. Throughout this process, we have ruled out many sentences as ungrammatical when they are acceptable in their non-literal interpretation. On the other hand, our system has allowed many sentences that are aspectually possible when we do not really use them. We are now in a position to say something about aspectual interpretation in context.

Before going on to the discussion of semantic/pragmatic clashes and reinterpretation problems, we have one important distinction to which we should draw attention: the distinction between the (tacit) knowledge of language structure and (tacit) knowledge of language use. Chomsky has made this distinction explicit with the introduction of generative linguistics. The following is the standard Chomskyan view of this distinction.

> What is knowledge of language? Answer: language is a computational system, a rule system of some sort. Knowledge of language is knowledge of this rule system.
>
> How is language used? Answer: the use of language is rule-governed behavior. Rules form mental representations, which enter into our speaking and understanding. A sentence is parsed and understood by a systematic search through the rule system of the language in question.
>
> (Chomsky 1988: 24)

He makes clear distinction between the system of knowledge of language (i.e. linguistic structure) and the various processing systems that put such knowledge into use. Following Chomsky in this respect, let me summarize the difference between the theories of language structure and its use.[20] Here, since our information on such processing systems are so scarce, I confine myself to utterance interpretation. The theory of grammar generates all the grammatical structures and rules out all the ungrammatical structures. In this sense, the grammar is just a source of ambiguities, and does not provide the best interpretation for utterances. Furthermore, sentences in isolation do not provide enough information for utterance interpretation either. The theory of utterance interpretation, on the other hand, is deterministic, and uses more information than language provides. In this sense, utterance interpretation employs abductive and goal-oriented procedures, where contexts, memory, encyclopediac knowledge, reasoning ability, etc. are involved. The grammar may be considered as instructions and constraints on utterance interpretation.

### Relevance Theory: A Theory of Utterance Interpretation

A good candidate for the theory of utterance interpretation is Relevance Theory proposed by Sperber and Wilson (1986).

> Linguistic decoding provides input to the inferential phase of comprehension; inferential comprehension involves the construction and manipulation of mental representations. An utterance can thus be expected to encode two basic types of information: representational and computational, or descriptive and procedural - that is, information about the representations to be manipulated, and information about how to manipulate them. (Wilson and Sperber 1993: 2)

If we take the relevance-theoretic point of view, the goal of communication is to attain maximal cognitive effects with minimal processing efforts:

(47) PRINCIPLE OF RELEVANCE:
Any utterance addressed to someone automatically conveys the presumption of its own relevance.

(48) RELEVANCE:
a. Other things being equal, the greater the cognitive effect achieved by the processing of a given piece of information, the greater its relevance for the individual who processes it.
b. Other things being equal, the greater the effort involved in the processing of a given piece of information, the smaller its relevance for the individual who processes it.
(Wilson and Sperber 1988: 140)

In this theory, idealized participants of conversation cooperate with each other to make their conversation optimally relevant. The speaker guarantees the utterance to be as easy as possible for the hearer to process, and the hearer makes the best hypothesis about the speaker's intention, by picking up a context that makes the relevance maximal. Technically, the best interpretation is the one that causes most effective information change in the hearer's database with least processing effort, or alternatively, the one that implies maximal cognitive effect with minimal deductive effort. In this sense, utterance interpretation is not bottom-up: the hearer's task in utterance interpretation is (i) to form an explicature of utterance that leads to a relevant implicature and (ii) to form relevant implicature based on the explicature of the utterance. The explicature formation procedure assumed in this study is sketched in Figure 4 below.

The best interpretation among alternatives: With this relevance-theoretic view in mind, let us go back to discuss aspectual interpretations. First, consider the following sentences again, repeated here as:

(49) #John killed the old man twice.

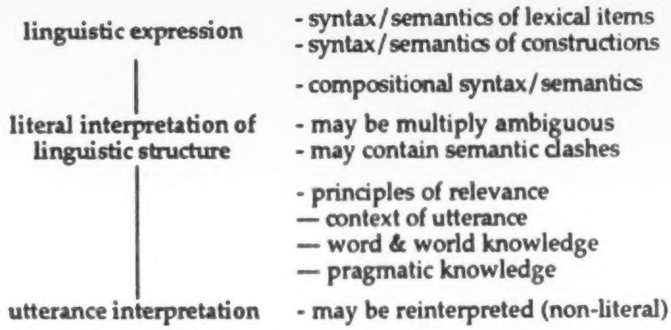

linguistic expression
- syntax/semantics of lexical items
- syntax/semantics of constructions
- compositional syntax/semantics
literal interpretation of linguistic structure
- may be multiply ambiguous
- may contain semantic clashes
- principles of relevance
— context of utterance
— word & world knowledge
— pragmatic knowledge
utterance interpretation
- may be reinterpreted (non-literal)

Figure 4: Formation of Explicature

Our aspectual semantics judges (49) as well-formed, even though we do not usually use such sentences. Remember that even if semantics permits unavailable interpretations, as it might, they simply will not be available. Parsability is not a requirement that must be met by the semantics. But, why does this sentence seem uninterpretable? Because the repetitive interpretation in (49) contradicts our belief about the world. Under the assumption of relevance, (49) would be interpreted as a metaphor (John did something very horrible to an old man twice), or the hearer would ask the speaker for more information to make (49) relevant. However, when the hearer has the information activated in his/her database that we cannot die twice, (for example, when s/he is considering whether human beings can be revived), the sentence can be relevant, because (49) in this context leads the hearer to infer that the speaker is saying that human beings can actually revive with (49) as its evidence.

Next, let us consider the paradigm in (50) and (51) again. Remember we treated linguistic expressions in (50) and (51) the same in that both denote an achievement event. Since achievements are incompatible with the progressive, the only grammatical option for sentences in (50)-(51) is to be interpreted as the progressive of iterative activity.

(50) a. Jenny was knocking at the door.
b. Bill is blinking now.
c. Mary is hitting John.

(51) a. The patient was dying.
b. John is arriving in five minutes.
c. Mary was reaching the top.

This iterative reading is, however, ruled out for (51), because it contradicts our normal assumption of the world that such events as dying and John's arriving do not repeat in regular situations. Hence, in this case, literal interpretation cannot be relevant. Since utterance interpretation relies heavily on the speaker/hearer's common sense knowledge as well as the context of utterance, when the speaker wants to mean literal but non-salient interpretation, s/he needs to provide enough context to make his/her utterance relevant to the hearer (i.e. easily accessible to the hearer), or to use some other means to convey the same information. When the literal reading is not salient, the hearer, on the other hand, needs to find some contextual effects that makes the utterance relevant. Thus, the sentences in (51) are reinterpreted as the progressive of an accomplishment. Note that this reinterpretation is blocked in (50) because their literal interpretation is relevant.

### Semantic Clashes and Reinterpretation

Let us now return to our main subject: reinterpretation.[21] I use the term 'reinterpretation' technically here to mean interpretation which is unavailable from grammar but attained by the principle of relevance. I do not claim that the hearer is actually re-interpreting the utterance in the face of semantic/pragmatic mismatches. Reinterpretation procedures are highly conventionalized and form a part of knowledge about the language. I distinguish three types of reinterpretation processes.

Shift of standpoint: The first type of reinterpretation is exemplified in (52). In (52a), the 'punctual' predicate *flash* is treated as referring to a durative event, and in (52b) the 'durative' predicate *climb the fire-ladder* is treated as denoting a punctual event.

(52) a. The light flashed for 10 milliseconds.
b. Moving along on the training course, she climbed the fire-ladder exactly at midday. (Talmy 1988)

The use of a durative or a punctual adverbial instructs the hearer to change his/her standpoint, in case of (52a) to zoom in, and in case of (52b) to zoom out. This shift of standpoint, or granularity of events in time, can be thought of as reconceptualization of situations in the world by moving from its 'default' standpoint to another. As Mittwoch (1980) puts it, 'like the geometrical point the punctual verb is an idealization'. This context dependent nature of event duration does not conflict with our thesis that there are four distinct eventuality categories, since what we are classifying is not the event in the world but the grammatical categories.[22] In this case, different ways of categorizing events tell us the relevant mode of speech by setting up the granularity level in describing events.

Accommodation of event concept: The second type of reinterpretation is illustrated in (53), where the 'durative' predicate *built a cabin* appears with a point-time adverbial. As Bach (1981) points out, (53) is possible in the context of Bill being a magician. More often than not, our normal assumption about the duration of events named by verbs is frozen, but here again the instantaneous adverbial guides our interpretation of the utterance. In this context, (53) conveys optimal relevance in that it makes the hearer infer that Bill can make a cabin instantaneously, as illustrated in (54). Note that the information in (54b)-(54c) cannot be expressed more economically by any other expression.

(53) At 4 o'clock, Bill built a cabin. (Bach 1981)

(54) a. Bill is a magician [premise]
b. Bill built a cabin at 4 o'clock [explicature of (53)]
c. Bill can make a cabin instantaneously. [implicature of (53)]

This reinterpretation process is different from the one discussed above. We do not change our standpoint but we change our assumption about what the cabin-building event is like. Reinterpretation of this type is not restricted to such a special case. Consider (55). One reading of (55) is John's opening the door is in progress at the time of utterance. In this reading, the use of the progressive leads us to interpret the door-opening event as durative.

(55) John is opening the door now.

This cannot be identified with the shift of standpoint, because (55) implicates that the event is taking more time than to be taken as an instantaneous event. It implicates that John is opening the door gradually perhaps because the door is heavy.

Metonymy link: The last and the most interesting type of reinterpretation involves metonymic associations. I assume that our encyclopedic knowledge includes event scripts, inference chains of situations metonymically (e.g. causally, temporally) related to particular events (cf. Schank and Kass 1988). The sentences in (56) are the first set of examples of this type of reinterpretation. In (56) the state expressions conflict with change-of-state adverbials.

(56) a. I knew *immediately* what he had in mind.
b. I am regretting it *already*. (Lyons 1977)
c. I will be back *in a minute*.

Under the presumption of relevance, the hearer introduces by inference an event metonymically related to the state named by the predicate, namely the event of coming into the state. Similar examples are given in (57).

(57) a. John needs $1000 *by Monday*. (ingressive reinterpretation)
b. I went to Tokyo *for three days*. (resultative reinterpretation)

This metonymical reinterpretation makes both of the conflicting expressions stay alive by introducing a new event by inference. In (56a), for example, the use of *immediately* tell the hearer that something happened immediately and the rest of the sentence tells him/her that the speaker knew what 'he' had in mind. The clashed predicate *know immediately* triggers metonymical association. The hearer infers that what happened immediately is the event of the speaker's getting to know what 'he' had in mind, based on his knowledge that if one gets to know something, s/he knows it. This inferential process is summarized as follows.

(58) Q: What happened immediately?
a. $\exists e$ Know(e) [linguistically introduced]
b. Come.to.know(e) $\rightarrow$ Know(e) [logical entailment]
c. $\exists e$ Come.to.know (e) [from a, b by inference]

This ingressive interpretation is triggered by crosssentential contexts as well:

(59) Max switched off the light. The room was pitch dark.
(Lascarides 1992)

We need to clarify one point here. This reinterpretation process is part of English speakers' knowledge about English. In other words, it is a conventionalized procedure of attaining non-literal interpretation. Yamanashi (1987) distinguishes three kinds of metonymical relations that characterize the associative processes of natural language:

(60) a. Semantic Metonymic Relation
b. Conventional Metonymic Relation
c. Pragmatic Metonymic Relation

This ingressive reinterpretation should be regarded as conventional metonymic process. First, this conventional aspect of this reinterpretation process becomes visible in crosslinguistic comparisons. Consider (61), Japanese counterpart of English sentences in (58).

(61) a. #Taroo-ga denki-o keshita. Heya-ga makkura-datta.
(lit.) 'Taro switched off the light. The room was pitch dark'
b. Taroo-ga denki-o keshita. Heya-ga makkura-ni natta.
'Taro switched off the light. The room became pitch dark'

In Japanese, this ingressive interpretation is unavailable for statives as shown in (61a). Japanese speakers need to use non-stative 'become' to convey the ingressive reading. This shows that the ingressive reinterpretation is not merely a reflection of general human cognitive ability, but rather a conventional procedure of utterance interpretation in English. However, this metonymic relation is still pragmatic and cannot be identified with a polysemic relation. This is illustrated in (62), where the ingressive reinterpretation is not readily available for the first discourse.

(62) a. #People began to leave. The room was empty. The janitors came in.
b. People began to leave. When the room was empty, the janitors came in.
(Partee 1984)

The ingressive reinterpretation is also available for activity expressions with point-time adverbials. Again, the point-time adverbial signals the hearer to find the relevant punctual event that is metonymically related to the durative event named by the verb, yielding an inchoative interpretation.[23]

(63) a. Mary ate dinner at six o'clock.
b. John ran at 4:30.

Accomplishments, however, do not allow this ingressive reinterpretation, as shown in (64). In other words, English does not have a conventional way of reinterpreting such sentences.[24]

(64) #John built a cabin at three o'clock.

The sentence (65a) below is an apparent counter example to this generalization. The inchoative interpretation of (65a) comes from the potential ambiguity of English indefinite article *a/an*. Since bare plurals do not express an indefinite quantity less than one, singular nouns with *the* or *a/an* sometimes make verb phrases ambiguous between activity and accomplishment, and the interpretation is variable according to the speaker and the context. The sentences in (65b) and (65c) support this view.

(65) a. John wrote a letter at 3:00.
b. #John wrote two letters at 3:00.
c. John wrote a/ #one letter for thirty minutes.

Temporal adverbials with *in* measure the period of time within which a change of state takes place, and the interval is measured from some pragmatically salient time-point. Consider the following utterance.

(66) Mary studied Japanese in an hour.

In a neutral context, this sentence would be interpreted as 'Mary started studying Japanese in an hour' via ingressive reinterpretation. In more contextual support, however, an activity-to-accomplishment reinterpretation is also possible through the accommodation of the event concept:

(67) I had a lot to do today. I studied Japanese in an hour, but the economics homework took forever to finish it.

In such case, the ingressive reinterpretation is ruled out by context. By reinterpreting *study Japanese* as an accomplishment of, say, finishing off the daily Japanese homework, the hearer can find a punctual metonym of the event named by the verb, namely the culminating point of the reinterpreted accomplishment.

Metonymical association and events in the progressive: Let us now extend the observation of metonymically motivated reinterpretation processes into progressive sentences. I assume that the function of the progressive operator in the event domain is to take a durative event and return a stage-level state where the event named by the verb is going on, which can be stated as follows (cf. Vlach 1981, Kawamura 1991a):

(68) FUNCTION OF THE PROGRESSIVE IN THE EVENT DOMAIN
   a. PROG is a function that maps a durational event into the corresponding stage-level state.
   b. Meaning postulate for the progressive
      If there is a state PROG($\phi$) that holds at time $t$, then there is an event $e'$ such that it is part of the event $\phi$ and $e'$ occurs at time $t'$ that precedes $t$.

Under this view, the punctual-durative contrast is the determinant of compatibility with the progressive: durative events can occur with the progressive, but achievements cannot.

As we can expect from the discussion above, achievements can appear in the progressive via reinterpretation, for example by shifting the standpoint:

(69) Fred was crossing the border at that moment. (zoom-in)

Reinterpretation in terms of metonymical association is also possible, and in fact provides more interesting examples.

(70) a. John is arriving.
   b. Mary was reaching the top.

In (70), the achievement-denoting expressions have conflicts with the progressive that requires a durative event. This category mismatch guides the hearer to look for a durative event that is metonymically related to the achievement named by the verb, based on his/her common sense about the world. Here, what seems to be happening is the introduction of a pragmatically determined preliminary stage that lead to the achievement event. This preliminary stage together with the achievement forms an accomplishment event, satisfying durative-event requirement by the progressive. Compare (70) with literal accomplishments in (71). The literal accomplishments do but the reinterpreted ones do not have the definite starting point of the process leading to the culmination. For example, in (70) we cannot be sure when John started arriving nor when Mary started reaching the top.

(71) a. Jenny is writing a paper.
b. Sean is studying for the final.

If I am on the right track, I can provide a new view to the so-called futurate progressives in English.[25] The following are the examples of the futurate construction:

(72) a. Mary is rehearsing tomorrow. (activity)
b. Jenny is writing a report next week. (accomplishment)
c. Lisa is reaching the moon at dawn. (achievement)
d. *John is knowing the answer tomorrow. (state)

This construction contains a semantic clash between the future-time adverbial and the progressive aspect, but is interpreted as referring to a 'preparatory state' of the event named by the verb (cf. Smith 1981, Dahl 1985, Kawamura 1991). Let us expand our analysis to this construction. First, consider (72c). This is identical to the examples in (70) with the difference that the former contains the future-time adverbial. In (70), we have literal achievement and reinterpreted accomplishment events. Let us assume that we have these in (72c) as well. The future-time adverbial guides the hearer to find an event that happens in the future. Fortunately, we have the literal accomplishment that is yet to happen at the time of utterance. Thus, the metonymic association can fix the semantic clash in (72c). But how can we make metonymically motivated the interpretation procedures for futurates in (72a) and (72b)? In these sentences, the literal events are already durative, being in perfect harmony with the progressive aspect. Given the interpretation procedure for achievement futurates, we can hypothesize that the use of future-time adverbial instructs us to adopt a zoom-out perspective. Be that as it may, the literally durative events as a whole are reconceptualized as punctual, making us to follow the same interpretation procedure for achievement futurates. Thus, futurate progressive constructions are analyzed as combinations of standpoint-shift and metonymy-link reinterpretations. Though it is subject to historical examination of the futurate construction, extension of the interpretation procedure from achievement progressive to futurate progressive in general through achievement futurates seems to be a plausible hypothesis.[26]

Before closing this section, I want to give an indirect support to my analysis of achievement progressives through its comparison with Japanese data. The following are the literal counterparts of English achievement progressives.

(73) a. Taroo-ga shin-deiru (Japanese)
Taro-NOM die-PROG-PRES
'Taro is dead / *?Taro is dying'
b. Hanako-ga ki-teiru
Hanako-NOM come-PROG-PRES
'Hanako has come / *?Hanako is coming'

As shown in the gloss, the reinterpreted accomplishment progressive reading is not available in Japanese. In Japanese, the progressive form is inherently ambiguous between on-going-process and result-state interpretations.[27] This indicates that the grammatically provided literal interpretation for achievement progressive preempts the accomplish reinterpretation in Japanese.

In this section, I have attempted to explain the reinterpretation phenomena based on the semantics proposed in the previous section, and also laid out the mechanism of utterance interpretation needed to make my claims precise. The proposed account for the reinterpretation procedures is motivated by human cognitive processes and the presumption of the existence of four eventuality types independently of verb meaning.

## Conclusion

Any adequate theory of aspect must answer at least the following three questions:

(Q1) What sort of aspectual classification is linguistically significant?
(Q2) How do extralinguistic factors enter into the aspectual semantics?
(Q3) What makes aspectual phenomena apparently so complex?

In this informal study, I have presented my partial and tentative answers to these questions. My answers to the first two questions are summarized as follows:

(A1) Classification of linguistic expressions and classification of semantic objects must be differentiated. As for the latter, I have argued for a four-way taxonomy on the line of Vendler (1957) in which the four categories have clear boundaries.
(A2) Linguistic information provides only partial information for the construction of the propositional meaning an utterance conveys.

> Different ways of categorizing situations in the world by language guide the hearer to make a best hypothesis about the meaning the speaker wants to convey.

These two answers together give us an answer to the third question. I would like to conclude:

(A3) The apparent complexity and subtleties of the aspectual phenomena found between verbs and sentences/utterances are the epiphenomena created by the interaction between grammar (aspectual semantics) and cognition (utterance interpretation).

## NOTES

* I wish to thank Manfred Krifka, Carlota S. Smith, and an anonymous KWPL reviewer for their comments and criticisms on earlier versions of this paper. Thanks are also due to Eric Dwyer and Anne-Marie Guerra for lending me their intuition about English. All errors and infelicities are, as always, my own.

[1] Following Bach (1981), I use the term 'eventuality' as a generic term for aspectual situation types.

[2] In the following tests, asterisks must be taken as indicating that the sentence in question is judged as ungrammatical in Vendler's grammar of aspect.

[3] By 'atomic sentences' I mean sentences that are yet to be connected with tense operators.

[4] I am not the first to state this idea explicitly; Vendler classes have been treated as a reference type of verbal predicates in algebraic semantics. See, for example, Link (1987) and Krifka (1989, 1992a) among others.

[5] The reason I am using only these particular tests among the rather varied ones in the literature is that these two tests are (i) sufficient for our present purposes, (ii) clearly temporal, so that they can be straightforwardly related to the temporal constitution of eventualities, and (iii) relatively free from atemporal factors that place further restrictions on their applicability. For a different view, see Smith (1991) among others.

[6] The terms 'accomplishment' and 'achievement' do not match our temporal characterization of eventualities they refer to. However, I follow the standard terminology even though they are misleading.

[7] The compatibility with the progressive can also serve as a test for discriminating durative and punctual events. The progressive will be discussed in Section 3.

[8] Whether a linguistic expression denotes a punctual or a durative event is often pragmatically determined, probably because it reflects how we perceive situations happening and holding in the world. See Section 3 for more discussion.

[9] Whether a durative expression denotes a homogeneous or heterogeneous event is grammatically determined by the predicate's thematic information.

[10] Pustejovsky (1991) considers that accomplishments consist of a process followed by a result state. Achievements are treated similarly as a transition decomposed into two states. The framework employed in this paper does not allow result states to be part of the event named by the verb. Rather, result states are treated as lexical entailment of the verb, since they are not in the extension of the predicate. I will discuss result states later in this section.

[11] Telic durative events are compatible with adverbials headed by *in*. *In*-adverbials measure the time span from some pragmatically determined time-point to the time-point where a change of state takes place. See Mittwoch (1980: 209-210) for more discussion on this point.

(i) a. *The kids played in three hours. (activity)
b. John ran a mile in three hours. (accomplishment)
c. Fred crossed the border in three hours. (achievement)

States and activities are incompatible with *in*-adverbials because the situations they denote are durative and contain no distinct transitional point in their temporal constitution. Being atomic events, achievements are compatible with telic *in*-phrases, because they are transitional points by themselves.

[12] For another view, see Verkuyl (1989) where activities are considered as having intermediate status between states and telic events.

[13] Individual-level state sentences are pragmatically odd with instantaneous adverbials. See Vlach (1981) for discussion. Note also that events

do have the durative interpretation when they are combined with the progressive or generic operator.

[14] Bach's (1989) grammaticality judgments are modified here to fit the characterization in the present theory, based on C. S. Smith's (p.c.) judgments.

[15] The GEN operator probably needs two arguments, a restrictor and a nuclear scope. See Krifka (1992b) for detailed discussion. For the distinction between individual-level and stage-level predicates, see G. Carlson (1977) among others.

[16] The need for multiple classification of linguistic expressions is also observed in Smith (1991).

[17] These verbs have yet another sense equivalent to *stand up* and *lie down*. We are not interested in such achievement readings here. For a different treatment of these verbs, see Smith (1991).

[18] The sentences in (37) have an activity reading as well, but do not discuss this reading because it is of irrelevance to the current discussion.

[19] Similar cases are found in nouns as well: *cake* may be either mass or count.

[20] I do not, however, identify knowledge of grammar with knowledge about language. Knowledge of language must include pragmatic knowledge as well as knowledge of grammar. See Gabbay and Kempson (1991) where the grammar is integrated into a framework of general reasoning.

[21] Moens and Steedman (1988) and Pustejovsky (1991a, 1991b, 1993) are brought to my attention after writing this paper. They employ the notion of 'coersion' to account for the phenomena described here as 'reinterpretation'. Pustejovsky defines the term as follows:

> TYPE COERSION:
> a semantic operation that converts an argument to the type which is expected by a function, where it would otherwise result in a type error. (Pustejovsky 1991b, 425)

Their views about the role of grammatical categorization in creative aspects of language use have much in common with the relevance-theoretic account advocated in this paper. However, I do not go into their proposals here in any detail.

[22] What eventuality a given linguistic expression 'normally' denotes may vary person to person, and utterance situation to utterance situation. The existence of variance does not weaken my claim. It is the values for variables that may be unstable and dependent on utterance contexts and individuals' normal assumptions about the world, but not event types themselves.

[23] When an activity event is completed, it must have both initial and final endpoints. We need to provide some cognitive motivation for the fact that only ingressive reinterpretation is available for such sentences, since this does not seem to be an arbitrary convention in English. A possible speculation would be that the initial endpoint is more accessible and hence relevant to the hearer because of its temporal precedence over the final endpoint, since our perception follows the temporal order of events.

[24] Probably this is because accomplishments have a heterogeneous internal structure. Accomplishments are a coordination of a process subevent and a culmination subevent, and the event introduced by ingressive reinterpretation is related only to the former subevent. However, the following sentence is perfectly ambiguous in Japanese.

(i) Watashi-wa 2-ji-ni uchi-e kaet-ta.
(lit) 'I returned home at two o'clock'
'I left (somewhere) for home at two o'clock'
'I went home and arrived there at two o'clock'

[25] Futurate progressives are discussed in Kawamura (1990) in more detail. See also Smith (1981, 1991) and references cited there.

[26] For conventionalization of pragmatic inferences, see Hopper and Traugott (1993).

[27] The Japanese progressive *te-iru* is morphologically decomposable into a verbal suffix *te* and a verb *iru* 'be/exist'. The progressive/result state ambiguity stems from the ambiguity inherent in *te*-linkage. See Hasegawa (1992) for extensive discussion on this topic.

## REFERENCES

BACH, EMMON. 1981. On time, tense and aspect: An essay in English metaphysics. Radical pragmatics, ed. by Peter Cole, 63-81. New York: Academic Press.

BACH, EMMON. 1986a. Algebra of events. Linguistics and Philosophy 9.5-16.

BACH, EMMON. 1986b. Natural language metaphysics. Logic, methodology and philosophy of science VII, ed. by Ruth Barcan Marcus, Georg J. W. Dorn, and Paul Weingartner, 573-595. Amsterdam: North-Holland.

BACH, EMMON. 1989. Informal lectures on formal semantics. New York: State University of New York Press.

CARLSON, GREG N. 1977. Reference to kinds in English. Amherst: University of Massachusetts dissertation. [Published, New York: Garland, 1980]

CARLSON, LAURI. 1981. Aspect and quantification. In Tedeschi and Zaenen, 31-64.

CHOMSKY, NOAM. 1988. On the nature, use and acquisition of language. Generative grammar: Its basis, development, and prospects, by Noam Chomsky, 1-32. Kyoto: Kyoto University of Foreign Studies. [Reprinted in Thirty years of linguistic evolution: Studies in honour of Rene Dirven on the occasion of his sixtieth birthday, ed. by Martin Pütz, 3-30. Philadelphia: John Benjamins, 1992]

DAHL, ÖSTEN. 1985. Tense-aspect systems. London: Blackwell.

DOWTY, DAVID. 1979. Word meaning and Montague grammar. Dordrecht: Reidel.

GABBAY, DOV M. and RUTH M. KEMPSON. 1992. Labelled abduction and relevance reasoning. SOAS Working Papers in Linguistics and Phonetics 2.41-84. London: School of Oriental and African Studies, University of London.

GAREY, HOWARD. 1957. Verbal aspect in French. Language 33.91-110.

GOLDBERG, ADELE E. 1992. Argument structure constructions. Berkeley: University of California dissertation.

HASEGAWA, YOKO. 1992. Syntax, semantics, and pragmatics of *te*-linkage in Japanese. Berkeley: University of California dissertation.

HINRICHS, ERHARD. 1985. A compositional semantics for aktionsarten and NP reference in English. Columbus: Ohio State University dissertation.

HOPPER, PAUL J. and ELIZABETH C. TRAUGOTT. 1993. Grammaticalization. Cambridge, UK: Cambridge University Press.

INAGAKI, DAISUKE and SEIJI IWATA. 1990. Aspectual classification of English verbs. Paper presented at the eleventh annual meeting of Tsukuba English Linguistics Society, University of Tsukuba.

KAWAMURA, MICHIHIKO. 1990. Eigo no futurate kobun (Futurate constructions in English). Tsukuba Eigo Kyoiku 11.17-31. Tsukuba Society of English Language Teaching.

KAWAMURA, MICHIHIKO. 1991a. On the semantics of the progressive: Toward an event-based alternative to interval semantics. Tsukuba Eigo Kyoiku 12.87-105. Tsukuba: Tsukuba Society of English Language Teaching.

KAWAMURA, MICHIHIKO. 1991b. How to represent time in English. Tsukuba: University of Tsukuba MA thesis.

KRIFKA, MANFRED. 1986. Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aspektklassen. München: Universität München dissertation. [Published, München: Fink, 1989]

KRIFKA, MANFRED. 1987. Nominal reference, temporal constitution: Towards a semantics of quantity. Proceedings of the sixth Amsterdam Colloquium, ed. by Jeroen Groenendijk Martin Stokhof, and Frank Veltman, 153-173. Amsterdam: Institute for Language, Logic, and Information, University of Amsterdam.

KRIFKA, MANFRED. 1989. Nominal reference, temporal constitution and quantification in event semantics. Semantics and contextual expressions. ed. by Renate Bartsch, Johan F. A. K. van Benthem, and Peter van Emde Boas, 75-115. Dordrecht: Foris.

KRIFKA, MANFRED. 1992a. Thematic relations as links between nominal reference and temporal constitution. Lexical matters, ed. by Ivan A. Sag, and Anna Szabolcsi, 99-123. Stanford: Center for the Study of Language and Information.

KRIFKA, MANFRED. 1992b. Genericity: An introduction. MS., University of Texas at Austin.

LAKOFF, GEORGE. 1987. Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

LASCARIDES, ALEX. 1992. Knowledge, causality, and temporal representation. Linguistics 30.941-973.

LEVIN, BETH. 1993. English verb classes and alternations: A preliminary investigation. Chicago: University of Chicago Press.

LINK, GODEHARD. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. Meaning, use, and interpretation of language, ed. by Rainer Bäuerle, Christopher Schwarze, and Arnim von Stechow, 302-323. Berlin: de Gruyter.

LINK, GODEHARD. 1987. Algebraic semantics for event structures. Proceedings of the sixth Amsterdam Colloquium, ed. by Jeroen Groenendijk Martin Stokhof, and Frank Veltman, 243-261. Amsterdam: Institute for Language, Logic, and Information, University of Amsterdam.

LYONS, JOHN. 1977. Semantics 2. Cambridge, UK: Cambridge University Press.

LYS, FRANZISKA, and KERRI MOMMER. 1986. The problem of aspectual verb classification: A two-level approach. CLS 21, Part 2: Papers from the parasession on pragmatics and grammatical theory at the twenty-second regional meeting, Chicago Linguistic Society, 216-230. Chicago: Department of Linguistics, University of Chicago.

MILLER, GEORGE A., and CHRISTIANE FELLBAUM. 1991. Semantic networks of English. Cognition 41.197-229.

MITTWOCH, ANITA. 1980. The Grammar of duration. Studies in Language 4.201-227.

MOENS, MARC, and STEEDMAN, MARK. 1988. Temporal ontology and temporal reference. Computational Linguistics 14.15-28.

MOURELATOS, ALEXANDER P. D. 1978. Events, processes, and states. Linguistics and Philosophy 2.415-434. [Reprinted with a few changes in Tedeschi and Zaenen, 191-212.]

PARTEE, BARBARA H. 1984. Nominal and temporal anaphora. Linguistics and Philosophy 7.243-286.

PLATZACK, CHRISTER. 1979. The semantic interpretation of aspect and aktionsarten: A study of internal time reference in Swedish. Dordrecht: Foris.

PUSTEJOVSKY, JAMES. 1991a. The syntax of event structure. Cognition 41.47-81.

PUSTEJOVSKY, JAMES. 1991b. The generative lexicon. Computational Linguistics 17.409-441.

PUSTEJOVSKY, JAMES. 1993. Type coercion and lexical selection. Semantics and the lexicon, ed. by James Pustejovsky, 73-94. Dordrecht: Kluwer.

QUINE, WILLARD V. 1960. Word and object. Cambridge, MA: MIT Press.

QUIRK, RANDOLPH, SIDNEY GREENBAUM, GEOFFREY LEECH, and JAN SVARTVIK. 1985. A comprehensive grammar of the English language. London: Longman.

SCHANK, ROGER, and ALEX KASS. 1988. Knowledge representation in people and machines. Meaning and mental representations, ed. by Umberto Eco, Marco Santanbrogio, and Patrizia Violi, 181-200. Bloomington: Indiana University Press.

SMITH, CARLOTA S. 1976. Present curiosity. Papers from the twelfth regional meeting, Chicago Linguistic Society, 568-581. Chicago: Department of Linguistics, University of Chicago.

SMITH, CARLOTA S. 1977. The vagueness of sentences in isolation. Papers from the thirteenth regional meeting, Chicago Linguistic Society, 568-577. Chicago: Department of Linguistics, University of Chicago.

SMITH, CARLOTA S. 1981. The futurate progressive: Not simply future + progressive. Papers from the seventeenth regional meeting, Chicago Linguistic Society, 369-382. Chicago: Department of Linguistics, University of Chicago.

SMITH, CARLOTA S. 1992. The parameter of aspect. Dordrecht: Kluwer.

SPERBER, DAN, and DEIRDRE WILSON. 1986. Relevance: Communication and cognition. Oxford: Blackwell.

TALMY, LEONARD. 1988. The relation of grammar to cognition. Topics in cognitive linguistics, ed. by Brygida Rudzka-Ostyn, 165-205. Amsterdam: John Benjamins.

TAYLOR, JOHN R. 1989. Linguistic categorization: Prototypes in linguistic theory. Oxford: Clarendon Press.

TEDESCHI, PHILIP, and ANNIE ZAENEN. 1981. Syntax and semantics 14: Tense and aspect. New York: Academic Press.

TENNY, CAROL L. 1986. Grammaticalizing aspect and affectedness. Masachusetts Institute of Technology dissertation.

VENDLER, ZENO. 1957. Verbs and times. The Philosophical Review 56.143-160. [Reprinted with minor changes in Linguistics in philosophy, by Zeno Vendler, 97-121. Ithaca, NY.: Cornell University Press, 1967.]

VERKUYL, HENK J. 1972. On the compositional nature of the aspects. Dordrecht: Reidel.

VERKUYL, HENK J. 1989. Aspectual classes and aspectual composition. Linguistics and Philosophy 12.39-94.

VLACH, FRANK. 1981. The semantics of the progressive. In Tedeschi and Zaenen, 271-292.

WILSON, DEIRDRE, and DAN SPERBER. 1988. Representation and relevance. Mental representations: The interface between language and reality, by Ruth M. Kempson, 133-153. Cambridge, UK: Cambridge University Press

WILSON, DEIRDRE, and DAN SPERBER. 1993. Linguistic form and relevance. Lingua 90.1-25.

YAMANASHI, MASA-AKI. 1987. Metonimic interpretation and associative processes in natural language. Language and artificial intelligence, ed. by Makoto Nagao, 77-86. North-Holland: Elsevier.

# NEGATIVE POLARITY ITEMS AND THE SEMANTICS OF THE PARTICLES *-TO* AND *-NA* IN KOREAN*

Yae-sheik Lee
*The University of Texas at Austin*

Abstract: The quantificational forces of polarity items are revisited and accounted for in terms of pragmatic principles and lattice theoretic semantics (cf. Krifka 1990). Particularly in Korean, the particles *-to* and *-na* are used to form polarity sensitive-*any* and free choice-*any* respectively. They are notorious for their various meanings. A unified analysis of them is pursued. Further I discuss what contribution they make to the semantics of polarity items. The paper also includes a brief discussion of the involvement of focus in polarity items.

## 1. Introduction

In this paper, I will treat four aspects of negative polarity items[1] (henceforth NPI) and the particles *-to* and *-na* in Korean. First, we are going to show what the particles *-to* , and - *na* mean. I will argue against previous analyses, which have assumed a variety of different meanings, and show that a unified analysis is possible. Second, based on the core meaning of *-to* and *-na*, which we are going to capture, I will provide an explanation for the reason that NPI's only go with the particle *-to*, and Free-Choice *amwu* ('any') with the particle *-na* . Third, I will briefly discuss the quantificational forces of polarity items, especially that of FC-*any*. Lastly, NPI's are accompanied by strong stress, arguably closely related to focus. Following Krifka (1992) we will try to show how the focus participates in the compositional derivations, and contributes to the entire meaning of the sentence in which it occurs.

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 89-123

## 2. The semantics of the particles *-to* and *-na*

Previous analyses: As mentioned above, most of the previous analyses of the semantics of *-to* and *-na* have argued or observed that they differ with the context in which they occur. We can classify the previous analyses into two groups: The first group are purely descriptive analyses and do not rely on any theoretical framework, and present intuitive meanings of *-to* and *-na* by taking into account the contexts in which the particles can occur. The second group are formal analyses. As representatives of the first and the second group of analyses, let's take Hong (1982 &1992 ), and Lee (1979), respectively. We will take his analysis of *-to* as representative because his analysis has been most influential, and is frequently referred to in the literature. Hong (1982) observes the following meanings of *-to* .

(1) Hong (1982:4 )

a. Picking-up an extreme case
Chomsky-*to* ku tongsa-lon-lul ihayhal-swu-ep-ta.
that syntactic-theory-Acc understand-can- not-Dec
'Not even Chomsky can understand the syntactic theory.'

b. Expressing the flavor of concession
Samtung cha-*to* coh-ta.
third-class train likable-Dec.
'The third class train is also OK'

c. Emphasizing negative assertion
.Mal ha-l him-*to* ep-ta.
talk do-Comp strength not have-Dec.
'(I) do not even have the strength to talk.'

d. Adding emphatic flavor
Eysang-*to* mos hay-ss-ta.
expectation cannot do-pst-Dec.
'(I) was not able to expect (what will happen.)'

e. Emphasizing the meaning of an adverb
Ppali-*to* talli-n-ta
fast run-prst-Dec
'Someone runs fast.'

Even though Hong (1982 & 1992 ) faithfully enumerates the meanings the particles *-to* and *-na* can convey, he doesn't try to explain why *-to* and *-na* change their meanings from context to context. If his observation were right, in other words, a word ( or morpheme) could have so many unrelated meanings, we would have to treat them as different lexical items, which happen to be phonologically identical. It is an implausible assumption that the particles *-to* and *-na* have more than one different meanings. Instead any theory that derives the contextually different meanings from a single core meaning and the influence of the context should be preferred.

Lee (1979), following Karttunen and Peters (1979), analyzes the meaning of an expression at two-levels, the meaning proper and the conventional implicature ( or presupposition ) of the expression, given as a pair < E, I >. He tries in particular to capture the meaning of implicature which seems to be triggered by the particles *-to* and *-na* Thus he observes that the particles *-to* and *-na* would have the following meanings, which are given here in informal paraphrases:

(2) Lee (1979:38 )
  a) Mica-to Chelswu-lul coahan-ta.
        -too        -Acc likes-Dec.
  E: Mica likes Chelswu.
  I: There is some x (x ≠Mica) such that x likes Chelswu.

  b) Mica-ka Chelswu-na manass-ta.
        -Nom        ?? met-Dec.
  E: Mica met Chelswu.
  I: i) There is some x (x≠Chelswu) such that Mica did not meet x.
    ii) Either Chelswu or x can equally be chosen for Mica to meet.
    iii) Neither Chelswu nor x is the best choice.
    iv) The best choice is not available ( for Chelswu to choose)

I think that his analysis of the meaning of *-to* and *-na* leaves many things to be desired. First, even though 2)-a) represents the implicature or presupposed meaning of the expression in which the particle *-to* appears, it alone cannot capture the meanings mentioned in (1) fully. In other words, Lee (1979) does not mention anything about the various meanings or how we can get such meanings relative to a given context. Second, he takes as the meaning of *-na* the meanings which seem to be mostly attributable to illocutionary operators. In short, he fails to take into account the fact that if the particle *-na* attaches to elements other than the subject, the sentences where it appears almost always are propositive but not declarative, as the following example shows.

(3)
a)*Mary-ka yenghwa-na po-ass-ta.
-Nom movie see-pst-Dec.
'Mary saw a movie'

b) Yenghwa-na po-ca
movie see-PRPS
'Let's see a movie'

In this paper, it will be argued that most of the implicature parts of *-na* in (2) are due to the propositive illocutionary operator.Third, Lee's framework inherits certain problems from Karttunen and Peters (1979), for example, the impossibility of variable binding across E and I, and the treatment of presupposition projection (cf. Beaver 1992 for a discussion). To avoid these problems and to allow for a compositional treatment of the derivations of expressions, we will introduce an operator for presupposition, '∂', following Beaver (1992). ( See Beaver 1992 for the semantics of this operator given in a framework of dynamic interpretation.)

An alternative analysis of the meaning of *-to* and *-na* : Following Bolinger (1977), we assume as the null hypothesis that an expression has only one (core) meaning. Other meanings are explained as being triggered by some pragmatic factors and our world knowledge or as arising through the interaction of the meaning of the particle with the meaning of its context in a compositional way. Keeping this in mind, let's turn first to the meaning of the particle *-to* . We can raise the question about how many meanings the particle *-to* could have. Hong's observations show that the meaning of *-to* varies with the context. However, a careful review of the examples in (1) tells us that all the uses of *-to* have something in common. This can be shown preliminarily as follows: Let us represent the meaning of a sentence minus the subject containing the particle *-to* by a predicate 'R' whose type is <e,t>. Every sentence asserts that R(x), where x is the denotation of the subject, and it is presupposed that there are other elements y that are comparable with x, for which R(y) holds. We see that by applying a standard test for assertion and presupposition, as mentioned in Van der Sandt (1988), the following hold: First, the presuppositions of a sentence F, but not its assertion, are entailed by a sentence like 'It is possible that F'. For example, the following shows the behaviors of presuppositions and assertions:

Test: It is possible that the king of France is bald.
⇒ There is a king of France. (Presupposition)
≠> The king of France is bald.(Assertion)

The second test to identify the presuppositional meaning of a certain sentence, is to make use of the following construction:

Test:
a) There is a king of France, and Mary regrets that the king of France is bald.
b) *Mary regrets that the king of France is bald, and there is a king of France.

The first part of a), 'There is a king of France.' is the presupposition of 'The king of France is bald.' If we utter b), it is infelicitous because the first part of b), 'Mary regrets that the king of France is bald.' has already presupposed the rest, 'There is a king of France.'. That is, this way of speaking violates the Cooperative Principle of Grice (1967) because the second conjunct is no longer informative to the hearer, having been assumed already by the first conjunct. The third test is to negate a sentence F that carries a presupposition, in which case the presupposition is entailed by the negation of F .

Test:
A: The king of France is not bald.
B: There is a king of France.

The third test is not so convincing in every case because with a negation the hearer can also reject and correct the speaker's presupposition. Thus the correction of the speaker's presupposition can look like a negation of the presupposition. By employing the first test[2], we can identify the presupposition of the speaker's utterances in (1) . I will do the test for (1)-a), b) and e)

(4) Tests:
a): Chomsky-to ku tongsalon-lul ihayhal swu epu-l swuiss-ta.
that syntactic theory-Acc understand can not-Rel can -Dec
'It is possible that even Chomsky cannot understand the Syntax.'

⇒ Chomsky-oy ihayhal swu-ep-nun salam-tul-i iss-ta.
-except understand can-not-Rel people-PL-Nom be-Dec.
'There are other persons than Chomsky, who can not understand it.'

≠> Chomsky-ka ihayhal-swu-ep ta.
-Nom understand-can-not Dec
'Chomsky can not understand it.

b). Samtung cha-to coaha-l swuiss-ta
third class train likable-Rel can -Dec..
' It is possible that the third class train is also OK'

⇒ Samtung cha-oy coaha-l cha-ka iss-ta.
third class train-except like-Rel train-Nom be-Dec.
'There are other trains than the third class train, which (I) like.

≠> Samtung cha-lul coaha-n-ta.
third class train-Acc like-prs-Dec.
'(I) like the third class train.'

e) Ppali-to tali-l swuiss-ta.
fast run-Rel can-Dec
'It is possible that (s)he runs also fast.'

=> Talun-pangsik-ulo-to tali-ess-ta
other-manner-in run-pst-Dec
'(S)he ran in other manners.'

≠> ppali tali-ess ta
fast run-pst Dec
'(S)he ran fast.'

These data suggest that we tentatively take the assertional meaning of *-to* to be a subject suffix; $\lambda X \lambda Y[Y(X)]$, that is, it simply says that the sentence without *-to* is true. As presuppositional meaning, we have $\lambda X \lambda Y \exists X' [X' \in ALT_{\neq}(X) \ \& \ Y(X')]$. Here X and Y are employed as variables over the element to which the particle *-to* attaches and the rest of the sentence in question, respectively. The notation 'ALT' is used to denote a function, which takes an object and yields a set of objects (of the same type as its argument) that may be regarded as being under consideration in an on-going discourse. '$ALT_{\neq}(X)$' stands for a set of alternatives to X which are not equal to X itself (cf. Rooth 1985 and Krifka 1991 for more on the use of the notation). That is, *-to* introduces the presupposition that there are alternatives to X for which the sentence predicate holds. If we represent the meaning as a two-level structure as above, we cannot guarantee that the X in the assertion part and the X in the presupposition part are identical. To guarantee this, we follow Beaver (1992), who introduces a presupposition operator '∂' within a framework of dynamic interpretation. Thus we can connect the two parts of meanings without the problem of insensitivity to variable bindings introduced in the assertion.

(5) $[\![\textit{-to}]\!] = \lambda X \lambda Y \, [\partial \exists X' \, [X' \in ALT_{\neq}(X) \ \& Y(X')] \ \& \ [Y(X)]]$.

Now let us explain why such a simple meaning of *-to* can give rise to seemingly different meanings in different contexts. For this purpose, we will refer to the concept of pragmatic scales as introduced by Fauconnier (1978). For example, (6)-a) pragmatically entails -b) according to (7).

(6)
a) Max can solve the most difficult problem.
b) Max can solve all problems .

(7) The pragmatic scale ≤;
≤ is a pragmatic scale for R ( predicate ) iff for all x, y in the domain of ≤, if R(x) and x ≤ y, then R(y) (Note that ≤ need not be linear or anti-symmetric, but for reasons of simplicity, here we assume that scales are linear.)

From the definition of the pragmatic scale we can derive the following principle, assuming that a speaker is always maximally informative:

(8)

A) If '$R(x)$' is uttered as the most informative expression among the alternative $R(y)$, $R(z)$,..., then we can conclude $\forall y[\, y \leq x \rightarrow \neg R(y)]$, otherwise $R(y)$ would have been uttered, with $y \leq x$ .

B) If '$\neg R(x)$' is uttered, then this is the most informative assertion among the possible alternatives. As $\forall x,y\ [\neg R(x)\ \&\ y \leq x \rightarrow \neg R(y)]$ ( from 7) ), this means that $\forall x,y\ [\, x \leq y \rightarrow \neg\neg R(y)]$

The pragmatic scale and the informativity principle above are closely related to Quantity Implicatures based on the maxim of Quantity by Grice (1967),'Make your contribution as informative as is required for the current purposes of the exchange.' Implicatures due to the maxim of Quantity have been analyzed by many scholars; Scalar Implicatures[3] by Horn ( 1972 ), Fauconnier (1975 &1978), Gazdar ( 1979 ), Levinson (1983), etc. What we should note is that scalar implicatures are built on the conventional elements of a linguistic scale in terms of Horn (1972); In (7), variables x, y belong to such a linguistic scale. In addition, (8), the informativity principle, shows the direction of quantity implicatures. Especially (8)-B) shows the classical cases of scale reversal by negation.

Let's turn to our data. What if we replace the negation, *ep* ('not have') in (1)-a) with *iss* ('be or exist')? All of a sudden, (1)-a) sounds unnatural or unacceptable. Why is this the case? It has to do with the pragmatic scale, (7) and the informativity principle (8). According to our[4] common world knowledge, the pragmatic scale would look as follows:

(9) Scale of understanding linguistic theories
< ...X, Y, Z... Chomsky >

According to (5), *-to* presupposes that there is at least one individual that is different from the referent of the NP to which *-to* attaches, and R, which in this case is **understand-the-theory**[5], holds of it. However, based on the pragmatic scale of (9) we know that if $\neg R$ ( **Chomsky** ), then

$\forall y[\ y \leq x\ \&\ x=\textbf{Chomsky} \rightarrow \neg R(y)]$. Or if 'Chomsky' doesn't understand the theory, then nobody can. This means that we can not find any individual x such that R(x) is felicitous as an assertion in terms of informativity. If the original negation is kept in situ, the negation reverses[6] the pragmatic scale according to (8)-B), and 'Chomsky' will be the bottom element on the reversed pragmatic scale. This reversal results in '$\forall y[\ x=\textbf{Chomsky}\ \&\neg R(x)\ \&\ x \leq y \rightarrow \neg R(y)]$.' Thereby, we can find individuals of which $\neg R$ holds. Therefore, the meaning of *-to* is compatible with the reversed scale. In short, to insure informativity the particle *-to* must not be attached to the top element on the pragmatic scale. Furthermore, the meaning of *-to* varies with the characteristics of the pragmatic scale concerned. For example, as for the flavor of concession in (1)-b), we can provide the following explanation;

(10)Scale of classes of trains ( in Korea)
< third class, second class, first class, prestigious class >

By the meaning of *-to* , and 'the third class of train', the denotation of the constituent to which the particle *-to* attaches, we know that there are other classes of trains, which are better than the third class, and naturally the speaker likes much more than the third class. Further we can infer that in order to express his modest concession the speaker picks up the bottom element on the scale of classes for trains, which is the most humble one available and goes to some extent against his own inclinations. On the other hands, if he utters 'the first class-*to* ' to show his yielding manner, he will fail to do so because on the scale of classes for trains, 'the prestigious class', which he can only yield by taking 'the first class' (the second luxurious one), is the most luxurious. Thus as a concessive clause, *iltung cha-to coh-ta* (' The first class train is also OK.') sounds awkward.

In conclusion, the role or the meaning of *-to* is only to indicate that there is some other option available on the pragmatic scale. or among alternatives We can relate this role to the notion of alternatives just by assuming that there are alternatives; if they forms a pragmatic scale, then the observed effects come about.

NPI's and *-to* : As in other languages, NPI's in Korean also are usually expressed by phrases denoting the minimal element of specified sorts[7]. However, as Krifka (1991:21) points out, 'it is often not easy to characterize a polarity sort, although one cannot help, as a speaker of a language, to have the idea that polarity items evoke a certain natural class of entities, events, attitudes, or the like.' The followings are examples of NPI's in Korean.

(11) a) Kunye-nun sonkkalak-to hana kkattak-ha-ci an-ass-ta.
she-Top finger one slight move-do-suff Neg-pst-Dec
'She didn't lift a finger.'

a')*Kuney-nun sonkkalak-to hana kkattak-hay-ss-ta.
she-Top finger one slight move-do--pst-Dec
'She lifted a finger.'

*sonkkalak-to hana kkattak-ha* ( 'lift a finger') relates to acts of labor, and denotes such an act that involves less labor than some arbitrarily small limit.

(11) b) Na-nun ttayngcen han-pwun-to ep-ta.
I-Top red cent one-counter(of money) not have-Dec
I don't have a red cent.'

b)*Na-nun ttayngcen han-pwun-to iss-ta.
I-Top red cent one-counter(of money) be-Dec
' I have a red cent.'

*ttayngcen han-pwun* ('a red cent') relates to amount of money, and denotes such an amount that is smaller than some arbitrarily small amount.

(11) c) Kuney-nun nwunsep-to hana kkattak ha-ci an-ass-ta.
she-Top eyelash one slight move do-suff Neg-pst-Dec
'She didn't bat an eyelash.'

c') *Kuney-nun nwunsep-to hana kkattak-hay-ss-ta.
-do-pst-Dec
'She batted an eyelash.'

*nwunsep-to hana kkattak-ha* ('bat an eyelash') relates to the reactions to stimuli, and denotes one that involves a weaker reaction than an arbitrarily weak reaction.

(11) d) Pi-ka han pangwul-to o-ci an-ass-ta.
rain-Nom one drop come-suff not-pst-Dec
'There was not a drop of rain.'

d') *Pi-ka han pangwul-to o-ass-ta.
'A drop of rain fell down.'

*han pangwul* ( 'a drop of (rain) ') relates to an amount of rain, and denotes such an amount that is smaller than some arbitrarily small amount.

(11) e) Ne-nun ku mwuncey-lul payknyen ka-to mos pwu-n-ta.
you-Top the problem-Acc 100 year go unable solve-pst-ta
' You can not solve the problem in 100 years.'

e') *Ne-nun ku mwuncey-lul payknyen ka-to pwu-n-ta.
' You can solve the problem in 100 years.'

In this case, the NPI, *payknyen ka* ( 'in 100 years' ) denotes very high values on the scale. Whether we select low or high values depends on the construction. Below I will explain why *payknyen ka* takes the highest value on the scale differently from most NPI's.

Following Fauconnier (1978), we know from the above examples that there is some ordering among the elements of a sort to which the NPI belongs. Krifka (1990 & 1991) makes use of a lattice-theoretic approach to explain the reason that NPI's have to appear in a (restricted) downward entailing (hereafter DE) context. Following him, we can assume that the elements of the same sort as the denotation of the NPI form a preorder but not a linear order as in Fauconnier's notion of the pragmatic scale., and that the NPI itself is the smallest element on that ordering.

(12) $S=\langle A', L_A, \leq_A\rangle$ is Negative Polarity Structure iff
a) if A' is of type $\delta$, $L_A$ is of type $\langle\delta, t\rangle$ ( set of $\delta$-elements);

b) $\leq_A$ is a preorder relation on $L_A$;
c) $A' \in L_A$, and $L_A$ contains more than one element;
d) $A'$ is the unique Y such that for every $X \in L_A$, $Y \leq_A X$.

For the above Negative Polarity Structure, we find a correlation either with the part relation or set-inclusion relation on events and objects as follows:

(13)
a) $\forall X, e\ [\ X \in L_A\ \&\ e \in X \rightarrow \exists e'\ [e' \in A'\ \&\ e' \leq_p e]$, where $\leq_p$ stands for a part relation.

b) $\forall X [X \in L_A\ \&\ X \neq A' \rightarrow X \subseteq A']$

For example, let X be some predicate applying to acts of labor, then it holds that there is an e' in the predicate **lift-a-finger** that is part of e. (13-b) is for the $L_A$ whose NPI, A' , is like **a-sound**, **anybody**, etc.. For (13-a), the part relation will turn into a set inclusion relation similar to (13-b) at the propositional level. This is shown in (13'). Since $\leq_A$ is defined in terms of a part relation for objects and events we can assume that the Negative Polarity Structure is compatible with the part relation for events and objects. (13) tells us that the set of possible worlds where an event or an object that the NPI applies to occurs or exists is the most inclusive, that is;

(13')
a) $\{\ w \mid [\![\ \exists e\ [\ \textbf{sing}\ (e)]\ ]\!]^{w}=1\} \subseteq \{\ w \mid [\![\exists e\ [\ \textbf{a-sound}\ (e)]\ ]\!]^{w}=1\}$

b) $\{w \mid [\![\exists e[\ \textbf{clean-a-room}(e)]]\!]^{w}=1\} \subseteq \{w \mid [\![\exists\ e[\ \textbf{lift-a-finger}\ (e)]]\!]^{w}=1\}$

Based on (13'-b), we can say that *sonkkalak-to hana kkattak-ha* ( 'lift a finger'), NPI, is the most inclusive one among the alternatives in $L_A$ . Therefore, it is the least informative alternative. Furthermore, we can generalize the infomativity relation based on the part or set-inclusion relation on objects or events as follows:

(13")$X \leq_I Y$ (Y is as much as or more informative than X), where X and Y belong to a set of same sort and are properties of t-based type, e.g., <s, <e,t>>, iff either

a)$\forall$ w', x, y[ X(x)(w') & Y(y)(w')&x $\leq_{p,w'}$ y $\rightarrow$[{w'| $[\![$Y(y)(w')$]\!]$ =1} $\subseteq$ {w'|$[\![$ X(x)(w')$]\!]$=1}]], or

b) $\forall$ w, X, Y [ Y(w) $\subseteq$ X (w)$\rightarrow$ {w | $[\![\exists$y [ Y(y)(w) ] $]\!]$=1} $\subseteq$ {w | $[\![$ $\exists$x [X(x)(w)] $]\!]$=1}]

Given the part relation and set inclusion and the informativity ordering above in connection with NPI, we can provide an explanation for the reason that NPI's appear in the contexts of DE. If NPI's are used in the scope of negation, or DE contexts, the set inclusion relation will be reversed, and this results in the reversal of the informativity-structure. Thus we have to use NPI's in the scope of negation or more generally in DE contexts because of the behavior of NPI's with regard to the relation between informativity and set-inclusion (or part relation), and the following pragmatic reason: If an utterance is an assertion, it has to meet the maxim of Quantity (Grice 1967), which says 'Make your contribution as informative as is required.' In particular: If alternatives are explicitly given, then a hearer can assume that the speaker chooses the most informative alternative. In other words, if a speaker makes an assertion A[X], where X belongs to a set of alternatives ALT(X), then (s)he must have reasons not to say A[Y], if $Y \in ALT(X)$. What do all these facts have to do with the occurrence of NPI's and the particle *-to* within the scope of negation? As shown in (5), the meaning of *-to* is related to the condition (12-c). Both express that the predication holds for an alternative. This is why NPI's and the particle *-to* can cooccur. The reason that their cooccurrence in an affirmative sentence is bad can be attributed to the same reason involving informativity as mentioned above. In the case of NPI's, the assertion part of the meaning of *-to* is a problem because the part $Y_{VP}(X_{NPI})$ is less informative than any other element in the polarity sort. For example, if for the sake of simplicity, we disregard the subject argument here and write **have** (P) for $\exists$x[ P(x)& **have**(x)], then (11)-b') roughly can be represented in our formalism as follows:

(14)[8]

$$[\,\partial\exists X'\ [X'\in \mathbf{ALT(red\text{-}cent)}\ \&\ \mathbf{have}\ (X')]\ \&\ \mathbf{have(red\text{-}cent)}]$$

The problem is that the assertion part of (14), **have(red-cent)**, is already presupposed by the presupposition part because $\exists X'[X'\in$ ALT **(red-cent)** & **have** (X')] always entails **have(red-cent)**, which is the assertion part of (14). Assuming that for an utterance to be a good (or felicitous) assertion, it has to be informative enough to shrink the input common ground. However, the utterance of (11-b') fails to shrink the input common ground because its assertion part is only the part of the presupposition. Therefore, it sounds awkward. As for (11)-e), Krifka (1989) nicely observes that the time-span adverbial (*payknyen ka* [9] 'in 100 years') plays the role of NPI. The explanation is as follows; If someone can solve a problem in two hours, then we can infer that (s)he can solve the problem in n hours where n is larger number than two. So we can generalize the inference pattern as follows; $\forall x, y\ \exists e\ [T(e) \subset_t y \,\&\, y \subset_t x \rightarrow T(e) \subset_t x]$, (where x and y are variables over time periods, e is a variable over events, T is a functor which takes an event to produce a time period, and $\subset_t$ is the relation of time-period inclusion.) The time-period of 100 years can be understood as a metaphor for a most inclusive time period belonging to the alternative set. Therefore, for *payknyen ka-to* to sound felicitous, it has to be employed in the scope of negation or DE context that reverses the ordering relation.

The semantics of *-na* and Free-Choice *amwu* ('any'): For the meaning of *-na* , first, let's look at cases where *-na* occurs with other than Free-Choice *amwu* ('any')( hereafter, FC-*amwu* ). In the same fashion in which we have tried to determine the presuppositional part and the assertional part of *-to* , we can apply our tests in the present case:

(15)

Chencay-na ku mwuncey-lul pwul swuiss-ta.
genius the problem-Acc solve can-Dec.
A genius can solve the problem.

Test: 'It is possible that ...'

Chencay-na ku mwuncey-lul pwul swuiss-nun kanungseng-i iss-ta.
genius the problem-Acc solve can-Rel. possibility-Nom be-Dec
'It is possible that a genius can solve the problem.'

⇒ Chencay-pota ttokhan salam-un pwul-swuiss-ta
genius - more than bright person-Top solve-can-Dec
'People brighter than a genius can solve it.'

≠> Chencay-nun pwul swuiss-ko ku pota ttokha-ci an-nun salam-nun pwul swu eps-ta.
-top solve can -and that brighter suff Neg-Rel -Top solve can not-Dec
'A genius can solve it and people less intelligent than a genius cannot.'

Based on the above test, (and the other tests which yield similar results) we conclude that an expression X-*na* has the following meaning :

(16) Presupposition : $\partial \forall P' [ P' \in ALT_{>}(P) \rightarrow \forall x [P'(x) \rightarrow Q(x)]]$.
Assertion: $\forall P'[ [ P' \in ALT_{<}(P) \rightarrow \forall x[P'(x) \rightarrow \neg Q(x)]] \& \forall x[P(x) \rightarrow Q(x)]]]$

(In (16), P stands for the variable over the element to which the particle *-na* attaches, and Q for the variable over the predicate, which is the rest of the sentence except the element with the particle *-na*. $ALT_{>}(P)$ and $ALT_{<}(P)$ ( an abbreviational notation) denote the alternatives to P that are a higher and lower value of the scale, whatever scale it may be, respectively. )

To put together what we have developed so far, the meaning representation of *-na* can be given as follows:

(17) $[\![ \textit{-na} ]\!] = \lambda X \lambda Y[ \partial \forall P''[ P'' \in ALT_{>}( X) \rightarrow Y(P'' )] \&$
$\forall P'[ P' \in ALT_{<}(X) \rightarrow \neg Y ( P')] \& Y(X)]$,
where X is for the variable over the element to which the particle *-na* attaches, and Y for the variable over the predicate, the rest of the sentence except X.

To illustrate (17), let's represent (15) as in (18) in a formal way.

(15) Chencay-na ku mwuncey-lul pwul swuiss-ta.
genius the problem-Acc solve can-Dec.
'A genius can solve the problem.

(18) $=> [\ \partial \forall P''[\ P'' \in ALT_{>}(\text{genius},) \rightarrow csp(\ P'')]\&$
$\forall P'[P' \in ALT_{<}(\text{genius}) \rightarrow \neg\ csp(P')]\&\ csp(\text{genius})]$, where **csp** stands for the predicate **can-solve-the-problem**.

This means that every alternative to a genius that is less intelligent than a genius cannot solve the problem but a genius can solve the problem. In addition, it is presupposed that those alternatives who are more intelligent than a genius can solve the problem. This representation appears to bear the meaning of (18) quite faithfully.

In Korean, as with *any* in English, people have distinguished Free Choice(FC)-*amwu* ('any') and Polarity Sensitive (PS)-*amwu* ('any'), which appears to have an existential quantifying force. As expected, PS-*amwu* goes with the particle *-to* in a negative sentence just like any other NPI. However, FC-*amwu* can occur with the particle *-na* . One thing to note is that the particle *-na* is usually accompanied by a modal auxiliary *swuiss* ('can'). In the literature, there have been many arguments about the quantifying force of *any* in English (of universal or existential). Below I will try to show why the FC-*amwu* and the PC-*amwu* seem to bear on the universal quantifying force and the existential quantifying force respectively in the framework of lattice theoretic approach. We can say that *amwu* in Korean denotes the most general term in the sort of persons, but actually it is different from the general term 'a person' in the sense that *amwu* is an NPI, and thereby introduces an ordered set of alternatives. Its lattice sort is defined as a set which contains every subproperty of 'person'. We can represent the meaning of (19) by using (17) as follows;

(19)

| | | | |
|---|---|---|---|
| Amwu-na | ku paty-ey | o-l | swuiss-ess-ta. |
| anyone | the party-at | come-Rel | can-pst-Dec. |

'Anyone can come to the party.'

$=> [\partial \forall P''[\ P'' \in ALT_{>}(\mathbf{amwu}) \rightarrow ccp(\ P'')]\&$
$\forall P'[\ P' \in ALT_{<}(\mathbf{amwu}\ ) \rightarrow \neg ccp(P'\ )\ ]\ \&\ ccp(\mathbf{amwu})]$ ,
where **ccp** is an abbreviation for the predicate **can-come-to-the-party**, and **amwu** stands for the meaning of **person**.

The presuppositional part means that every individual which belongs to every subproperties of person can come to the party. This is true, given the assumption **ccp ( amwu )**[10]. The first conjunct of assertional part is also vacuously true because there is no property in ALT ( **amwu** ) that could be more inclusive than **amwu**. Even if it is vacuously true, it contribute to the emphatic assertional force of the entire sentence together with the second conjunct because it tells us that there are no persons that can not 'come to the party.' Notice that according to (12)-d), **amwu** is the unique bottom element of that polarity structure, which includes all the subproperties. Hence we can also infer that FC *amwu -na* has the universal quantifying force. For the existential quantifying force of PS-*amwu -to* is discussed in the next section.

It is remarkable that when *-na* attaches to expressions other than the subject, the sentences are not declarative. They are usually propositive similar to the form of 'Let's' in English.

| | | | |
|---|---|---|---|
| (3) | a)*Mary-ka | yenghwa-na | po-ass-ta. |
| | -Nom | movie | see-pst-Dec. |
| | 'Mary saw a movie' | | |
| | b) Yenghwa-na | po-ca | |
| | movie | see-PRPS | |
| | 'Let's see a movie' | | |

Below we will explain how *-na* can have the meaning of the implicature part from the (core) meaning of (17) *-na* (as in Lee 1979) of (2)). So far we have seen what the core meanings of *-to* and *-na* look like, and why *-to* has to appear with NPI's in negative sentences, while *-na* has to go with FC *amwu*. Until now, we excluded from our discussion the focus marked by the stress on NPI's to which *-to* , and *-na* attach. In the next section, we are going to deal with focus, illocutionary operators, and the matter of compositional derivations of the predicted readings.

## 3. Compositional derivation of representations with focus

Focus and *-to* and *-na* : The problem of focus is so involved that we cannot deal with it in detail in this paper. However, in this section we are going to talk briefly about the representation of focus and about ways of interpreting focus that serve our current purpose. There are various ways of realizing focus. Syntactically it is realized by cleft constructions, rightward movement, etc. To put a stress on a certain element in a sentence can also be a way of marking focus. As we mentioned above, NPI's to which *-to* or *-na* attach bear a strong stress. and this indicates that this expression is in focus. When we want to derive the meaning of a sentence with a focus compositionally, we have to devise a way to project the information about where the focus is located from the focus expression itself to the complex expression. For the interpretation of focus, two frameworks are popular; namely, Structured Semantic Representations and Alternative Semantics. The second was developed by Rooth (1985) to treat focus-sensitive operators such as *only* . According to this theory, a semantic representation in focus is related to a set of alternatives. As for the first framework, the presupposition analysis is prevalent. As representatives, we can take Jackendoff (1972), and Williams (1980), where a sentence with the focused part is analyzed to be partitioned into two parts a presupposition set and a focus or Focus-Presupposition structure, '< F, P>'. This idea has been developed to the concept of Focus-Background structured meaning, '<B, F>' by von Stechow (1989), Jacobs (1990) and Krifka (1992). In short, focused elements will introduce structured meanings, and the Focus-Background structures are interpreted by focus operators, e.g., illocutionary operators like ASSERT(ion), QUEST(ion), DIRECT(ive), OPTAT(ive). For example the meaning of ASSERT(<B, F>) as follows;

(20) ASSERT(<B,F>): It is asserted that B(F), with the felicity condition that the values X such that B(X) are under discussion.

(20') [JOHN]$_F$ came to the party.

According to (20), the utterance of (20') asserts that B(F), in other words, 'John came to the party' with the following felicitous condition that the hearer and speaker have been discussing who came to the party. The felicity conditions vary with different Focus-Background structures. To derive the representation of sentences with focus compositionally, we need some compositional rules for recursive structured meanings. In other words, we have to devise a rule which enables the focus-background information to be conveyed to higher nodes. Following Krifka (1992), we are going to make use of the following recursive definition of application (21).

(21) ( Krifka 1992:.25)

a) If $\alpha$ is of type $(\delta)\tau$ and $\beta$ is of type $\delta$, then $\beta(\alpha)$ is of type $\tau$ and is interpreted as functional application.

b) Focus inheritance from operator;
if $<\alpha,\beta>$ is of type $<(\delta)(\tau)\mu,\delta'>$, and$\gamma$ is of type $\tau$, then $<\alpha,\beta>(\gamma)$ is of type $<(\delta)\mu, \delta'>$, and is interpreted as $<\lambda X_{\delta}[\alpha(X)(\gamma)], \beta>$.

c) Focus inheritance from argument:
If $\gamma$ is of type $(\delta)\tau$ and $<\alpha,\beta>$ is of type $<(\mu)\delta, \mu'>$, then $\gamma(<\alpha,\beta>)$ is of type $<(\mu)\tau,\mu'>$, and is interpreted as $<\lambda X_{\mu}.[\gamma(\alpha(X))], \beta>$.

In addition, in this paper, we will make use of the following notations: x, y, z, x', z', etc. as variables of type of e. P, P', etc. are variables of type <e,t>, and a, b are meta-variables . If A is a syntactic form, then $[\![A]\!]$ is the semantic representation of A in our semantic representation language. The following rule covers focusation of a constituent; if a constituent is focused then it introduces a background-focus structure; the background is empty at the point, that is, a simple identity function.

(22) SF:C $\rightarrow$ $C_F$(arbitrary category C indexed by focus feature F);
$[C_F] = <\lambda X.X, [C]>$, where X is of the type from which the type of [C] is derived that is not a focus-background type.

Thanks to type-raising theories such as those of Partee and Rooth (1983) and Dowty(1986), we can generalize the meaning representations of *-to* , (6), and *-na* , (19) as follows to allow for argument categories of various type.

(23)[11]
a)$[\![$*-na*$]\!]$: $\lambda\alpha\ \lambda\beta\ \lambda\nu\ [\ \partial\forall\alpha''[\alpha''\in ALT_{>}(\alpha) \rightarrow [\alpha'', \beta](\nu)]\ \&$
$\forall\alpha'[\alpha'\in ALT_{<}(\alpha) \rightarrow \neg[\alpha',\beta]]\ \&\ [\alpha,\beta](\nu)]$

b) $[\![\textit{-to}\,]\!]$: $\lambda\alpha\lambda\beta\,\lambda v\,[\,\partial\exists\alpha'\,[\alpha'\in \mathrm{ALT}_{\neq}(\alpha)\ \&\ [\alpha',\beta](v)]\&\ [\alpha,\beta](v)]$

Derivations and interpretations: First, let's show a simple compositional derivation of (24),which does not contain focused elements;

(24) Chencay-na ku mwuncey-lul pwul swuiss-ta.
genius the problem-Acc solve can-Dec.
'A genius can solve the problem.

*chencay* (genius); **genius**: type; <e,t>
|
| *-na* :$[\![\textit{-na}\,]\!]$: NA, the abbreviation of the semantics of *-na*.
| /
*chencay-na* :NA( **genius**)
| $= \lambda\alpha\,\lambda\beta\,[\,\partial\forall\alpha''[\,\alpha''\in \mathrm{ALT}_{>}(a)\rightarrow\beta(\alpha'')]\&$
| $\forall\alpha'[\alpha'\in \mathrm{ALT}_{<}(\alpha)\rightarrow\neg\beta(\alpha')]\ \&\ \beta(\alpha)]](\textbf{genius})$
|
| $=\lambda\beta\,[\,\partial\forall\alpha''[\alpha''\in \mathrm{ALT}_{>}(\textbf{genius})\rightarrow\beta(\alpha'')]\&$
| $\forall\alpha'[\alpha'\in \mathrm{ALT}_{<}(\textbf{genius})\rightarrow\neg\beta(\alpha')]\ \&\ \beta(\textbf{genius})]]$
| *ku mwunchay-lul pwul swu iss* .
| / $=\lambda P\forall x[P(x)\rightarrow \mathbf{csp}(x)]$[12]
*Chencay-na ku mwunchay-lul pwul swu iss* :
$=\partial\forall\alpha''[\alpha''\in \mathrm{ALT}_{>}(\textbf{genius})\rightarrow\forall x\,[\alpha''(x)\rightarrow\mathbf{csp}(x)]]\&$
$\forall\alpha'[\alpha'\in \mathrm{ALT}_{<}(\textbf{genius})\rightarrow\neg\forall x\,[\alpha'(x)\rightarrow \mathbf{csp}(x)]]\&\forall x[\textbf{genius}(x)\rightarrow$
$\mathbf{csp}(x)]$

To provide a full interpretation for the sentences like (24), which don't contain focused element(s), we need the following assertional illocutionary operator;

(24') ASSERT ($\phi$), where $\phi$ is a variable over propositions, shrinks a common ground c to a common ground c',where c' is the intersection with the set of possible worlds for which $\phi$ is true. Felicity conditions; (i) $c'\neq c$ (asserting $\phi$ makes a difference in the common ground), (ii) $c'\neq 0$ (the truth of must not be already excluded by c)

If we apply the above assertional operator, ASSERT, to the derivational result of (24), we get the following:

| *-ta* : ASSERT
|/
*Chencay-na ku mwunchay-lul pwul swu iss ta* :
$= \mathrm{ASSERT}(\partial \forall \alpha''[\alpha'' \in \mathrm{ALT}_{>}(\textbf{ genius}) \rightarrow \forall x\ [\alpha''(x) \rightarrow \textbf{csp}(x)] \&$
$\forall \alpha'[\alpha' \in\ \mathrm{ALT}_{<}(\textbf{ genius}) \rightarrow \neg \forall x\ [\alpha'(x) \rightarrow \textbf{ csp}(x)] \&$
$\forall x[\textbf{ genius}(x) \rightarrow \textbf{ csp}(x)]\ )$

The assertion of (24) changes the common ground to those worlds where it is presupposed that every kind of person that ranks higher than a genius on a scale of intelligence can solve the problem, and every kind of person that ranks lower than a genius on a scale of intelligence cannot solve the problem. Of course, every person who is a genius can solve the problem in those worlds.

Next, to deal with sentences like (3) we also need an propositive illocutionary operators, (e.g., PROPOS), which is defined as follows;

(25)[13] PROPOS(< B, F>) = If there is no X, X∈ ALT(F) such that to make B(X) is preferred to the speaker and the hearer over B(F), then the speaker propose to make B(F) true.

let's try to derive the meaning representation of (3), which is a propositive sentence.

(26)
*yenghwa* : $=\lambda R \lambda x \exists y[\textbf{movie}(y) \& R(x,y)]$ =**M** for the abbreviation of
| the representation of *yenghwa*, as an object NP
SF [*yenghwa* $]_F$ : $<\lambda TT, M>$
|
| *-na* : ⟦*-na* ⟧:NA, the abbreviation of the representation of *-na*.
|/
*yenghwa-na* :

| $= \lambda\alpha\ \lambda\beta\lambda v\ [\ \partial\forall\alpha''[\alpha''\in ALT_{>}(\alpha) \rightarrow [\alpha'',\beta](v)]\&$
| $\forall\alpha'[\alpha'\in ALT_{<}(\alpha)\rightarrow \neg[\alpha',\beta](v)]\&[\alpha',\beta](v)](<\lambda TT, M>)$
|
| $=<\lambda T'\ \lambda\beta\lambda v\ [\ \partial\forall\alpha''[\alpha''\in ALT_{>}(T') \rightarrow [\alpha'',\beta](v)]\&$
| $\forall\alpha'[\alpha'\in ALT_{<}(T')\rightarrow \neg[\alpha',\beta](v)]\ \&[T',\beta](v)]],\ M>$
| *po* :**see**
|/
*yenghwa-na po* :
| $=<\lambda T''\ \lambda v\ [\ \partial\forall\alpha''[\alpha''\in ALT_{>}(T'') \rightarrow [\alpha'',\mathbf{see}](v)]\&$
| $\forall\alpha'[\alpha'\in ALT_{<}(T'')\rightarrow \neg[\alpha',\mathbf{see}](v)]\ \&[T'',\mathbf{see}](v)]],\ M>$
| *wuli* (we) : λPP(**we**)
|/
*wuli yenghwa-na po* :
| $=\lambda PP(\mathbf{we})(<\lambda T''\ \lambda x[\ \partial\forall\alpha''[\alpha''\in ALT_{>}(T'') \rightarrow[\alpha'',\mathbf{see}](x)]\&$
| $\forall\alpha'[\alpha'\in ALT_{<}(T'')\rightarrow \neg[\alpha',\mathbf{see}](x)]\&[T'',\mathbf{see}](x)]],\ M>)$
|
| $=<\lambda T'''[\ \partial\forall\alpha''[\alpha''\in ALT_{>}(T''') \rightarrow \alpha''(\mathbf{see})(\mathbf{we})]\&$
| $\forall\alpha'[\alpha'\in ALT_{<}(T''')\rightarrow\neg[\alpha'(\mathbf{see})(\mathbf{we})]\ \&\ T'''(\mathbf{see})(\mathbf{we})]]\ ,\ M>$
|
| *-ca* ; PROPOS
|/
*eynghwa-na po-ca* :
$=PROPOS(<\lambda T'''[\ \partial\forall\alpha''[\alpha''\in ALT_{>}(T''')\rightarrow\alpha''(\mathbf{see})(\mathbf{we})\&$
$\forall\alpha'[\alpha'\in ALT_{<}(T''')\rightarrow\neg[\alpha'(\mathbf{see})(\mathbf{we})]\ \&\ [T'''(\mathbf{see})\ (\mathbf{we})]]\ ,$
$\lambda R\lambda x\exists y[\mathbf{movie}(y)\&R(x,y)]\ >)$

We get the following representation:

(27)
If there is no X, X∈ ALT(**movie**) such that to make true **See**(X), is preferred to the speaker and the hearer over **See(movie)**, then the speaker propose to make **See(movie)** true.

The above result tells us everything that Lee (1979) chose to express for *-na* in (3). In other words, since **See** (X) and **See** (**movie**) are abbreviations for B(X) and B(F), respectively they can be fully spelled as follows;

(28)

(i) $B(F){:}\partial\forall\alpha''[\alpha''\in ALT_{>}(\mathbf{movie})\rightarrow\exists y[\alpha''(y)\ \&\ \mathbf{see}\ (\mathbf{we}, y)]]\&$
$\forall\alpha'[\alpha'\in ALT_{<}(\mathbf{movie})\rightarrow\neg[\exists y[\alpha'(y)\ \&\ \mathbf{see}\ (\mathbf{we}, y)]\ \&$
$\exists y[\mathbf{movie}\ (y)\ \&\ \mathbf{see}\ (\mathbf{we}, y)].$

(ii) B(X):similarly, we can get the fully spelled B(F) by replacing movie in (i) with X .

From (27) and (28), we can say that propositive (26) means that the speaker proposes to the hearer that whatever they (alternatives to a movie) may be, if they rank lower on a scale of entertainment (e.g., less exciting) than a movie does, then the speaker and hearer see a movie instead, while if there is an alternative which is more preferable to the speaker and the hearer, they see it. In addition, what is noteworthy with propositives containing the particle *-na* is that they convey the politeness of the speaker in the following sense: the speaker can always leave some room for the hearer to make a counter-proposal by employing the particle *-na* because the presupposition part triggered by the particle *-na* says ' if 'you', the hearer, have or know something better than what 'I', the speaker, propose, then let us do it.'. In the case of (26), the movie is in fact the best option that is allowed for them, but there is an implicational meaning that the speaker doesn't take the proposal as the really best one to convey her or his politeness. But such a reading is not directly triggered by the meaning of the particle *-na*, but by the conspiracy of the illocutionary operator and *-na* .

Let us now turn to the case of NPI, *amwu*, with the particle *-na* . Following Krifka (1992), we apply an illocutionary operator ASSERT to the sentence radical to obtain the meaning of a sentence with the declarative mood marker. Let's modify (20) to some extent to serve our purpose better by specifying the felicity condition more precisely. Following Stalnaker (1979), we assume that an assertion modifies or shrinks the shared assumption or common ground, which can be represented as a set of possible worlds, and a sentence S can also can be represented as a set of possible worlds where S holds.

(29)

ASSERT(<B, F>) shrinks a common ground c to a new common ground c', such that $c'=c \cap [\ B(F)]$. Felicity conditions required;

i) Asserting B(F) results in a different common ground.,$c \neq c'$.
ii) The result of $c \cap [\ B(F)]$ is not an empty set.
iii)There is an X which is an element of ALT(F), and B(X) could have been asserted with regard to c that would have a different result. That is $c \cap [\ B(X)]$ is not an empty set $c \cap [B(X)] \neq c$, and $c \cap [\ B(F)] \neq c \cap [\ B(X)]$.

Next, let's see why (30)-a) sounds awkward, but -b) is ok in the course of the derivation.

(30)

a)*Amwu-to ku pati-ey wa-ss-ta.
anyone the party-at come-pst-Dec.
*'anyone came to the party.'

b) Amwu-to ku pati-ey o-ci an-ass-ta.
anyone the party-at come-suff Neg-pst-Dec.
'Nobody came to the party

*amwu* (any); $\lambda P\exists x[\mathbf{amwu}(x)\&P(x)]$;type; $<<e,t>,t>$
|
SF *[amwu]*$_F$ : $< \lambda TT,\lambda P\exists x[amwu(x)\&P(x) >$; T is of type $<<e,t>,t>$
|
| *-to* : TO:the abbreviation of the sematics of *-to*.
|/
*amwu-to* ;
| $=TO(<\lambda TT,\lambda P\exists x[\mathbf{amwu}(x)\&P(x)]>)$
|
| $=<\lambda T.\ \lambda\beta\ [\ \partial\ \exists\alpha'\ [\alpha'\in ALT_{\neq}(\ T')[\alpha',\beta]]\&[T',\beta]],$
| $\lambda P\exists x[\mathbf{amwu}(x)\&P(x)]>$
|
| wa-ss('came'); **came**
|/
*amwu-to wa-ss*
| $=<\lambda T.\ \lambda\beta\ [\ \partial\ \exists\alpha'\ [\alpha'\in ALT_{\neq}(\ T')[\alpha',\beta]]\&[T',\beta]],$
| $\lambda P\exists x[\mathbf{amwu}(x)\&P(x)]>(\mathbf{came})$
|
| $=<\lambda T''[\partial\exists\alpha'[\alpha'\in ALT_{\neq}(T'')\&[\alpha'(\mathbf{came})]]\&[T''(\mathbf{came}]],$
| $\lambda P\exists x[\mathbf{amwu}(x)HP(x)]>$
| *-ta* ; ASSERT;

I /
*Amwu-to wa-ss-ta* :
$=\mathrm{ASSERT}(<\lambda T''[\partial\exists\alpha'[\alpha'\in \mathrm{ALT}_{\neq}(T'')\&[\alpha'(\mathbf{came})]]\&$
$[T''(\mathbf{came})]],\lambda P\exists x[\mathbf{amwu}(x)\&P(x)]>)$

According to (29), to be an appropriate assertion, the above should meet the following conditions;

(31)

(i) B(F),that is, $\partial\exists\alpha'[\alpha'\in \mathrm{ALT}_{\neq}(\lambda P\exists x[\mathbf{amwu}(x)\&P(x)])\&$ $\alpha'(\mathbf{came})]\ \&\ \exists x[\mathbf{amwu}(x)\&\mathbf{came}(x)]]$, is meaningful and informative enough to result in a new (different from the input common ground) common ground.

(ii) it is under discussion for which $X\in\ \mathrm{ALT}(\lambda P\exists x\ [\mathbf{amwu}(x)\&\ P(x)])$, it holds that B(X),that is $\partial\exists\alpha'[\alpha'\in \mathrm{ALT}_{\neq}(X)\&\alpha'(\mathbf{came})]\&X(\mathbf{came})]$ would have a different result from that of $B(\lambda P\exists x[\mathbf{amwu}(x)\&P(x))$.

(iii) the speaker has reasons not to assert B(X), $X\in \mathrm{ALT}_{\neq}(\lambda P\exists x[\mathbf{amwu}(x)\&P(x)])$

Taking into account the fact that the focused part is an NPI, *amwu* ('any'), which is the unique bottom element of the polariy structure, we can say that if there is any X, which is different from *amwu* ('anyone') and hence belongs to a more specific set than '*amwu* ('anyone') came to the party.', then it is automatically asserted that a person came. In other words, the presuppositional part entails what the assertional part means. This means the above utterance is under no circumstances informative enough to change the input common ground to a new different common ground. In this sense, the above utterance fails to meet the conditions of (31), and therefore, it sounds awkward. As for (30)-b), it differs from (30)-a) in that it is a negative sentence. Thus the following felicitous conditions should be met for (30)-b) to be a good assertion.

(31')

(i) B(F),that is, $\partial\exists\alpha'[\alpha'\in \mathrm{ALT}_{\neq}(\lambda P\exists x[\mathbf{amwu}(x)\&P(x)])\&$

$\neg\alpha'(\mathbf{come})] \ \& \ [\neg\exists x[\mathbf{amwu}(x)\&\mathbf{come}(x)]]$, is meaningful and informative enough to result in a new (different from the input common ground) common ground.

(ii) it is under discussion for which $X\in ALT(\lambda P\exists x[\mathbf{amwu}(x)\& P(x)])$,it holds that $B(X)$,that is $\partial\exists\alpha'[\alpha'\in ALT_{\neq}(X)\&\neg\alpha'(\mathbf{come})]\& [\neg X(\mathbf{come})]$ would have a different result from that of $B(\lambda P\exists x[\mathbf{amwu}(x) \ \& \ P(x))$.

(iii)the speaker has reasons not to assert $B(X)$, $X\in ALT_{\neq}(\lambda P\exists x[\mathbf{amwu}(x)\&P(x)])$

If we assume that $\neg\exists x[\mathbf{amwu}(x)\&\mathbf{come}(x)]$ is true, then the presuppositional part is also obviously true because of the following reasons. The presuppositional part says that individuals of a proper subproperty of **person** in general didn't come to the party. While the assertional part says that every individual x which **amwu** applies didn't come.This means that individuals belonging to any subproperty of **person** in general didn't came since any individuals, which belong to subproperties of **amwu** are also elements of **amwu**. The set of possible worlds where the assertional part is true is a proper subset of The set of possible worlds in which presuppositional part is true due to negation, which is complementation on sets of worlds, thereby reverses set inclusion. Therefore, if the assertional part is true, then the presuppositional part is also true. This means that the first felicitous condition of (31') is met. The assertional part is informative enough to reduce the input common ground to a new common ground. As for the second and third conditions of (31'), the speaker has good reasons not to pick up an individual that holds for a subproperty of **person** in general. For example, if the speaker says, ' a 'beautiful lady' didn't came to the party.', this utterance would result in a weaker assertion than that bearing *amwu* because the set of possible worlds where 'a beautiful lady didn't came to the party' is a superset of the set of the possible worlds of the asssertion that 'nobody came to the party.' Thus the speaker employs *amwu* to make a strong assertion instead of other expressions denoting a subproperty of **person** in general. Therefore (30)-b) is a good assertion and the meaning of the particle *-to* rightly serves to make the sentence a felicitous assertion. Obviously, this way of interpreting *amwu-to* used in a negative context leads to the reading of a negation of the existential quantifier.

As for adverbials with the particle *-to* , we can also show their derivation in the same manner as above. The derivation is as follows;

(32) ppali-to wa-ass-ta
fast come-pst-Dec
(S)he came fast (as well as in other manners)

*ppali* (fast); **fast** , whose type is <<e,t>,<e,t>>
|
SF [*ppali* ]$_F$: < λEE, **fast**>; E is of type <<e,t>,<e,t>>
| *-to* :TO:the abbreviation of the sematics of *-to*.
| /
*ppali-to* :
| =TO(<λEE, **fast**>)
|
| =λ α λβ λx [ ∂∃α'[α'∈ ALT$_{\neq}$( α)&[α',β](x)&
| [α,β](x) ](<λEE,**fast**>)
|
| =<λE'.λβλx [∂∃α'[α'∈ ALT$_{\neq}$( E'.)&[ α',β](x)[E'.,β](x)],**fast**>
|
| *wa-ss* ('came'); =λy[**came**(y)]
| /
*ppali-to wa-ss:*
|
| =<λE'.λβλx [∂∃α'[α'∈ ALT$_{\neq}$( E'.)&[ α',β](x)&
| [E'.,β](x)],**fast**> (λy[**came**(y)])
|
| =<λE".λx [∂∃α'[α'∈ ALT$_{\neq}$(E".)&[ α',λy[**came**(y)]](x)&
| [E",λy[**came**(y)](x)] ,**fast**>
|
| somebody; λPP(c(x)); c(x) means discourse context fixes who x is
| /
(Somebody)*ppali-to wa-ss*
| =λPP(c(x))(<λE".λx [ ∂∃α'[α'∈ ALT$_{\neq}$(E".)& [α', λy[**came**(y)]](x) &
| [E",λy[**came**(y)](x)] ,**fast**> )
|
| =<λE".[ ∂∃α'[α'∈ ALT$_{\neq}$(E") &[ α',**came**(c(x))] &
| [E",**came**(c(x))], ,**fast**>
| *-ta* ; ASSERT;

| /
(Somebody) *ppali-to wa-ss-ta*;
$=ASSERT(\lambda E''[\partial\exists\alpha'[\alpha'\in ALT_{\neq}(E'')\&\alpha'(\mathbf{came}(c(x)))\&$
$E''(\mathbf{came}(c(x)))]], \mathbf{fast}>)$

According to (29), we can obtain the following representation from (32).:

(33)
$ASSERT(\lambda E''[\partial\exists\alpha'[\alpha'\in ALT_{\neq}(E'')\&\alpha'(\mathbf{came}(c(x)))]\&$
$E''(\mathbf{came}(c(x)))]], \mathbf{fast}>)$ shrinks a common ground c to a new common ground c', such that $c'=c \cap$
$[\partial\exists\alpha'[\alpha'\in ALT_{\neq}(\mathbf{fast}) \& \alpha'(\mathbf{came}(c(x)))] \& \mathbf{fast}(\mathbf{came}(c(x)))]$.

Felicity conditions required;
i) Asserting $[\partial\exists\alpha'[\alpha'\in ALT_{\neq}(\mathbf{fast})\& \alpha'(\mathbf{came}(c(x)))]\& \mathbf{fast}$ $(\mathbf{came}(c(x)))$ results in a different common ground, $c \neq c'$.

ii) The result of $c \cap [\partial\exists\alpha'[\alpha'\in ALT_{\neq}(\mathbf{fast}) \& \alpha'(\mathbf{came}(c(x)))]\&$ $\mathbf{fast}(\mathbf{came}(c(x)))]$ is not an empty set.

iii)There is an X which is an element of ALT(**fast**), and B(X) could have been asserted with regard to c that would have a different result. That is $c \cap [B(X)]$ is not an empty set, $c \cap [B(X)] \neq c$, and $c \cap [B(\mathbf{fast})] \neq c \cap [B(X)]$.

According to (33), for (32) to be a good assertion, first, $[\mathbf{fast}(\mathbf{came}(c(x)))]$ must be informative enough to shrink the input common ground c, and $[\exists\alpha'[\alpha'\in ALT_{\neq}(\mathbf{fast})\&\alpha'(\mathbf{came}(c(x)))]]$ must be presupposed in the common ground, which means that someone came fast as well as in other manners. Second, the truth of $\partial\exists\alpha'[\alpha'\in ALT_{\neq}(\mathbf{fast})\&\alpha'(\mathbf{came}(c(x)))]\& \mathbf{fast}(\mathbf{came}(c(x)))]$ must not be already excluded by the input common ground c. In other words, the speaker must not assert anything that has been already taken as false. Third, the speaker and the hearer have discussed in what manner (s)he came, and similarly, '(s)he came in a different manner from 'fast' (e.g.., *safely* )' could have been felicitously asserted with regard to the input common ground c. Such an assertion result in a different common ground from the one in which the assertion ' (s)he came *fast* .' results. However, the speaker has a good reason not to make an assertion with other alternatives to *fast*, e.g., *safely*. One reason we could imagine is that the speaker would like to make an assertion on the fastness of his coming rather than on the *safeness* of his coming. This is why Hong (1982) observes that the particle *-to* also has the function of emphasizing the meaning of an adverb[14]. Actually this is due to the illocutionary operator, ASSERT.

## 4. Concluding Remarks

As mentioned in the introduction , we have accomplished three tasks; We showed that the particles *-to* and *-na* can be analyzed as having one core meaning with their other accompanying meanings due to illocutionary forces or to our world-knowledge on orderings among the elements of alternative sets. Second, we argued that both the incompatibility of *amwu* ('any') plus *-to* and the coocurrence of *amwu* ( 'any') plus *-na* with upward entailing property predicates are due to informativity and some other pragmatic felicity conditions. Third, the quantifying force of PS-*amwu* (∃) and FC-*amwu* (∀) are due to the conspiracy of the meanings of the particles *-to* and *-na*, ordered alternatives introduced by the *amwu* and the status of *amwu* in the ordered alternatives, and pragmatic principles (e.g., scalar implicature ) Last, following Krifka (1992), we tried to derive sentences where 'alternative sensitive' operator *-na* and *-to* introduce a strong stress as focus on the element to which they attach. The compositional derivations done so far show what the representations of meanings of the expressions with focus and illocutionary operators look like. In addition, we argued that Lee (1979)'s analysis of *-na* , and *-to* leaves many things to be desired, and most of the problems of Lee (1979) can be solved by making use of illocutionary operators as in Krifka (1991, and 1992). In this paper we leave how to treat the GEN(eric) operator undetermined for further study.

## NOTES

* I thank Byung-Que Jun, Han-chang Kim, Kee-chan Kim, Robert Wall , Manfred Krifka, and Jack Hoeksema for their comments and suggestions. I

especially thank Manfred Krifka for his direction of this paper. I also benefited from insightful discussions of ideas in this paper with all members of *the Semantics Club* in the University of Texas at Austin . All errors, of course, are mine.

1 Here we mean by NPIs a set of expressions that typically appear in specific contexts, most prominently the scope of negation, and denote the minimal value or small entity of the sort concerned, or general things, etc. Therefore they are named Negative Polarity Items.

2 With the second and third tests, the same result as that in (4) will be obtained. For the purpose of saving space, I do not do the tests here.

3 Levinson(1983) presents the following rule for deriving scalar implicatures from scalar predicates. Given the scale $< e_1, e_2, \ldots e_n >$, if a speaker asserts A(e2), then he implicates $\neg A(e_1)$, if he asserts $A(e_3)$, he implicates $\neg A(e_2)$ and $\neg A(e_1)$, and in generally if he asserts A(en), then he implicates from $\neg A(e_{n-1})$ up to $\neg A(e_1)$.

4 Here 'our' means those who know something about linguistics and especially current syntactic theories. Thus by 'our' knowledge, we know that Chomsky has the most amount of knowledge of current syntactic theories and is the most likely to understand the syntactic theories.

5 A word in boldface stands for the representation of its meaning .

6 The reversing effect by negation is due to Fauconnier(1975 &1978).

7 In Heim(1984), remedying Ladusaw (1979)'s general Downward Entailing Contexts for the occurrence of NPI's, which can not explain some contexts like if-clauses, she restricts the admissible strengthening to those which are introduced by the alternative items in the position of the NPI.

8 When it related to the lattice sort, the alternative set ALT is interchangeable with the lattice sort, $L_A$.

9 For Korean, one hundred might have been the largest number imaginable. Actually *ka* in Korean means 'to go', but in this case, its meaning is corresponding to the meaning of 'in' in English.

10 Actually **amwu** and its alternatives can be taken as generalized quantifiers so that they take verbal predicates as their argument: **Amwu** ( **ccp** ). For simplicity, here I take them as if they are of type e.

11 $[\alpha, \beta]$ means that either of them can be a functor or an argument of the other. We can always predict this thanks to the type theory . We make use of '$\lambda v\ [\alpha, \beta](v)$' to deal with the elements which are other than the subject, e.g., the object. $v$ is a variable over a sequence of arbitrary number of variables so that if $[\alpha, \beta]$ is a one place-predicate, then v means one variable x. While if $[\alpha, \beta]$ is a sentence, than $v$ means the zero number of variables.

12 Here the universal quantifying force is due to the hidden Generic operator, which is usually interpreted to have more or less as universal quantifying force. But it is better to recognize the fact that universal quantifier and GEN(eric) operator are different in the following sense that while GEN allows for exceptions, the universal quantifier does not allow for exceptions. I have to leave this problem for further study since I have no idea on how to remedy this flaw.

13 This is just a preliminary sketch of the propositive illocutionary operator. I would like to leave elaboration of illocutionary operator in formalism and illocutionary operators in general, for a further study.

14 In English, often the meaning of 'also' acquires the meaning of 'even'. In terminology of Koeing (1991), both are 'additive' particles . Hoeksema, J. comments that the particle *-to* should be translated into 'even' in English since most of the contexts in which the particle *-to* seems

to appear correspond to those which 'even' in English appears in. I think that the particle *-to* in Korean fills both functions of 'even' and of 'too' or 'also' because *-to* can attaches to a constituent that doesn't introduce any scale as follows:

Tom -to wa-ss-ta
come-pst-Dec
' Tom came, too.'

## REFERENCES

Beaver, D. 1992. "The Kinematics of Presupposition." *ITLI Prepublications for Locic, Semantics and Philosophy of Language.*

Bolinger, D. 1977. *Meaning and Form* . Longman.

Dowty, D. 1986. "Type raising, fumctional composition, and non-constituent conjunction." *Categorial Grammar and Natural Language Structure*, ed by R. Oerler, E. Bach, &D. Wheeler. Dordrecht: Reidel, 153-198.

Fauconnier,G. 1975. "Pragmatic Scales and Logical Structure." *Linguistic Inquiry 4*, 353-375.

Fauconnier,G.1978. "Implication Reversal in a Natural Language." *Formal Semantics and Pragmatics for Natural Languages* ed. by F. Guenthner and S.J. Schmidt, 289-301.

Gazdar, G. 1979. *Pragmatics: Implicature, Presupposition and Logical Form.* York: Academic Press.

Grice, P. 1967. "Logic and Conversation William James Lecture, Harvard University. Partly published as 'Logic and Conversation'." *Syntax and Semantics 3: Speech Act* ed.(in 1975) by P.Cole & J.L. Morgan Academic Press, New York. 41-58.

Heim,I. 1987. "A Note on Negative Polarity and Downward Entailingness". *NELS 14*, 98-107.

Hong, S-M. 1982. "A Comparative Study of Korean and Japanese Bound-Morphemes(III) :With special reference to the Semantic analysis of ' Pole-Meaning Delimiters." *Kyungpopok-National-University Language Research 16*. 67-94.

Hong, S-M. 1992. *A Study of Lexical Semantics of Korean*. Hakmwunsa, Seoul.

Horn, L.R. 1972. *On the semantic properties of Logical Operators in English.* Ph. D. Diss. UCLA.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar.* Cambridge, Mass.: MIT Press.

Jacobs, J. (to appear). "Focus ambiquity."In: *Preceedings of Conference on Focus and Intonation, Fraundhofer Institut Stuttgant*, eb by J. Hoepelman.

Karttunen, L. & S. Peters. 1979: "Conventional Implicature." *Syntax and Semantics 11: Presuppositions* ed by Ch.-K. Oh &D.A. Dinneen.New York; Academic Press. 1-56.

Keoning, E. 1991. *The Meaning of Focus Particles-A Comparative Perspective*. Routledge. London and New York.

Krifka, M. 1989. "Nominal Reference, Temporal Constitution and Quantification in Event Semantics." *Semantics and Contextual Expressions* ed. by R. Bartsch, J.van Benthem, and P. vanEmde Boas. 75-115.

Krifka, M.1990. "Polarity Phenomena and Alternative Semantics. " *Preceedings of the 7th Amsterdam Colloquium, Amsterdam, ITLI* ed. by M. Stokhof and L. Tolenvliet. 277-302.

Krifka, M.1991: "Some Remarks on Polarity Items." Ms. Dep't. of Linguistics, Univ. of Taxas at Austin.

Krifka, M.1992. "A Compositional Semantics for Multiple Focus Construction." *Information-structure and Grammar*. ed. by J. Jacobs et al.

Ladusaw, W. 1979. *Polarity Sensitivity as Inherent Scope Relations*, Ph.D. Dissertation, University of Taxas at Austin.

Lee, I-H. 1979. *Korean Particles, Complements,and Questions; A Montague Grammar Approach*, .Dissertation,University of Texas at Austin.

Levinson, S. C. 1983. *Pragmatics*. New York: Cambridge Univ. Press.

Partee, B.& M. Rooth. 1983. "Generalized Conjunction and Type Ambiquity." *Meaning, Use, and Interpretation of Language*, ed. by R. Bauerle, et al.

Rooth, M. 1985. *Association with Focus*, Dissertation, Univ. of Mass., Amerst.

Stalnaker , R. 1979. "Assertion." *Syntax and Semantics 9.-Pragmatics*, ed. by P. Cole. New York: Academic Press. 315-332.

Stechow, A. v. 1989. "Focusing and Backgrunding Operators." Univeriat Konstanz, Fachgruppe Sprachwissenschaft, Arbeitspapier Nr. 6. Konstanz.

Van der Sandt, R.A. 1988. *Context and Presupposition*. London: Croom Helm.

Williams, E. 1980. "Remarks on Stress and Anaphora "*Journal on Linquistic Research* 1; 71-82.

# Clitics, Case Checking, And Causative Constructions

Xavier Villalba
Universitat de Barcelona

Abstract: In this paper it is argued that the distribution of third person pronominal clitics in Catalan Causative Constructions may receive a principled account if two independently motivated processes are assumed: Det Incorporation, a proper means to check Case, and Verb Incorporation, a process at stake in Causative Constructions. This analysis also gives a proper answer to the theoretical problems that optional Climbing Climbing poses for Chomsky's (1992) Last Resort Principle.

The distribution of (third person) pronominal clitics in Catalan causative constructions (hereafter: CC) raises many puzzling questions that have not received a principled and comprehensive analysis. Consider the well-known contrast between external and internal arguments of the embedded verb with respect to Clitic Climbing (hereafter: ClCl):

(1) a. *Faré dormir-los junts.
(I) make-FUT sleep-them(ACC) together
b. Els faré dormir junts.
(I) them(ACC) make-FUT sleep together
'I will make them sleep together.'
(2) a. Faré posar-los junts a la Maria.
(I) make-FUT put-them together to the Mary
b. Els faré posar junts a la Maria.
(I) them make-FUT put together to the Mary
'I will make Mary put them together.'

Whereas the causee clitic must climb, the object one may climb or not. It is tempting to account for this contrast in terms of an internal vs. external argument distinction. However, such a distinction is helpless if unaccusative verbs are taken into account:

(3) a. *Va fer venir-los aviat, els metges.
PAST make come-them(ACC) soon the doctors
b. Els va fer venir aviat, els metges.
them(ACC) PAST make come soon the doctors
'The doctors, he made them come soon.'

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 125-147

(4) a. Que va fer venir-ne molts, de metges?
that PAST make come-of.it many of doctors
b. Que en va fer venir molts, de metges?
that of.it PAST make come many of doctors
'Did he make many doctors come?'

The clitic related to the internal argument of the unaccusative verb venir 'come' is expected to behave as the object clitic in (2) with respect to ClCl: it should optionally climb. However, such an expectation is only partially fulfilled: when accusative, it must climb, just like causee clitics do; nevertheless, when partitive, it may climb or not, just like internal argument clitics do. An explanation on the basis of the internal vs. external nature of the argument must thus be revisited.

Furthermore, dative clitics pose another challenge because clitics related to dative internal arguments cannot climb, as the interpretative contrast between (5a') and (5a") shows:

(5) a. Li faré enviar[lzi] cartes.
him/her-DAT make-FUT send-them-DAT letters
a'. *'I will make them send letters to him/her.'
a". I will make him/her send them letters.'

The contrast between accusative and dative internal argument clitics is again an unexpected result when compared with the contrasts shown above.

After considering all these examples, it is extremely difficult to make any empirical generalization on the range of ClCl in Catalan CC. Moreover, the preceding examples raise many theoretical questions as well. Firstly, let us take into account the optionality of ClCl, as it is shown in (2). Chomsky's (1992) Last Resort Principle (LRP) states that operations (i.e., 'move a') take place just in order to make the derivation converge; in other words, if an operation need not take place, it cannot. Nevertheless, the existence of a phenomenon like optional ClCl prima facie poses a problem for the LRP. Another point is the relation between cliticization and Case. We must ask ourselves whether the different case of the clitic has any role in the distribution of clitics in Catalan CC, as seems to be the case. If the answer is affirmative, we must go further and investigate the place for cliticization in a theory of Feature Checking like that of Chomsky (1992).

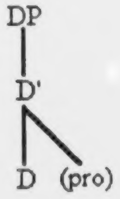
In this paper I will suggest that an analysis of CC and cliticization crucially involving head-movement (Verb Incorporation and Determiner Incorporation respectively) can account for the preceding contrasts. Moreover, it will be shown that such an analysis can also provide us with an answer for the two theoretical problems: the optionality of ClCl and the relation between cliticization and Case.

The organization of this paper is as follows. In the first section I will outline the analysis of cliticization I will assume throughout. Afterward, I will briefly work out the issues concerning CC. In the third section I will deal with the distribution of clitics in CC. Finally, the main conclusions will be addressed.

## Cliticization as Det Incorporation

In this paper it will be assumed, following Uriagereka (1989, 1992), Laenzlinger (1990) and Roca (1992), which pursued seminal unpublished work by E. Torrego, that cliticization is in fact Determiner Incorporation (hereafter: DI). This is not surprising if looked at from a diachronic perspective: both determiners and clitics have the same origin, Latin demonstratives. From a synchronic point of view, the above-mentioned authors have shown that there is evidence enough to assume that both determiners and clitics are Dets heading their own DP projection. Consequently, the representation I will assume throughout for clitics is the following:[1]

(6)



Given this analysis, cliticization is an instance of head movement. We thus expect its distribution to be constrained by general principles constraining movement (the ECP, if looked at from a standard approach, or a condition on chain formation, if looked at from a minimalist one). This paper is not the place to argue for the advantages of the 'Det Hypothesis' over the 'Affix Hypothesis'. Nevertheless, I will briefly show in the next paragraph how the former might subsume the basic intuitions of the latter, especially with respect to Case.

A Minimalist Approach to Cliticization. Borer (1984), the classical approach to the 'Affix Hypothesis' of cliticization, arrives at the following main conclusions: (a) clitics are generated as features of the head of their phrase -i.e., they are the 'spell-out' of the Case properties of the head; (b) the argument position is filled by a lexical or null NP depending on whether a Case-assigning device is available or not; (c) the clitic governs the argument NP; and (d) the clitic and the argument NP must be coindexed. Both (c) and (d) straightforwardly follow from the 'Det hypothesis': the moved Det must govern its trace and obviously they must have the same index. Conclusion (b) refers to the question of 'clitic doubling'. As Roca (1992) has argued, the 'Det Hypothesis' can account for this phenomenon without further stipulations (even though the version he adopts cannot): the doubled DP appears on the Spec of the DP headed by the clitic, the complement position being occupied by a pro. The status of this doubled DP is far from clear, but clarifying it would exceed the scope of this paper, so I am focusing my attention on question (a).

The intuition that clitics are the 'spell-out' of the Case properties of heads, even though productive, has not received a clear technical implementation in the literature. The approach to feature checking developed in Chomsky (1992) seems to provide us with a good solution. Nevertheless, some technical background is needed. The next paragraph will be devoted to sketching the basis of this approach. Afterward, we will deal with the relationship between cliticization and Case.

On Feature Checking. Chomsky's (1992) approach to the role played by morphological features in syntax is based on two main points: (a version of) the Strong Lexicalist Hypothesis (SLH) and a theory of feature checking. The SLH assumes, in the spirit of Lexical Phonology, that lexical elements are inserted in syntax with all their morphological features. Verbs will thus be drawn from the lexicon fully inflected for Tense and Agreement (V-features). The same holds for DPs with respect to Case and f-features (N-features). However, if feature assignment devices and, more important, the conditions regulating their application are dispensed with, How is the correct matching of features to be achieved? It is obvious that a theory of feature checking is needed.

Such a theory should give an answer to two basic questions: Why should features be checked? and How are features checked? Answering the first question is quite straightforward if the minimalist approach to linguistic theory is adopted. Chomsky (1992:6) claims that 'the linguistic expressions are the optimal realizations of the interface conditions, where 'optimality' is determined by the economy conditions of UG.' From this point of view, it seems quite plausible to regard morphological features as elements that may be dispensed with at LF, because they play no role at this interface level. Therefore, if a feature is not checked before LF-interpretation applies, the input to this interface level will not be optimal, making the derivation crash.

The question of how features are checked may be divided into two aspects: the formal checking device and the structural conditions regulating its application.

Regarding the first aspect, it is worth noting that while answering the question of why checking was necessary a sketch of a formal checking device has implicitly been suggested. If features must disappear before LF-interpretation applies, it is logical to regard checking as an elision process: features are checked, if they correctly match, they are eliminated; if they don't, they remain and the derivation crashes at LF.

The structural conditions regulating checking are quite straightforward: the (element carrying the) feature must be in a local relation with an appropriate head, i.e., a head carrying the same feature. Technically speaking, a feature must enter into the checking domain of a head carrying the same feature, where the 'checking domain' of a head a is the minimal set of nodes (the complement excluded) contained in MAX (a) that are distinct from a and do not contain a (see Chomsky 1992).[2] Namely, the checking domain of a head includes its specifier and anything adjoined (adjunction being allowed to its maximal projection, its specifier, or its head). A feature may be checked in any of these positions.

Case Checking. According to the SLH, argument DPs bear a Case when entering the syntactic component.[3] So then, they must enter into the checking domain of a Case-bearing head. Three logical options exist: raising to the specifier, adjunction to the specifier, and adjunction to the head.[4] The former -raising to the specifier- is, according to Chomsky & Lasnik (1991) and Chomsky (1992), the core realization of structural Case checking. Given the highly articulated structure of the clause, argument DPs, which are generated VP-internally, move to the specifier of their respective agreement heads (Agr-S and Agr-O). There they enter into the checking domain of V, which has previously adjoined to the Agr head to check its inflectional features (see Chomsky (1992) for a detailed analysis).

The second option -adjunction to the specifier- seems to be restricted to a single case: sentences with a there-DP pair of the kind:

(7) There is a unicorn in the garden.

The DP *a unicorn* bears nominative case, but it cannot check it in situ. Note that the presence of the expletive *there* makes raising to the specifier of Agr-S (or Tense) impossible. The only possibility is to adjoin the DP to the expletive in the specifier of Agr-S/Tense (at LF), checking its nominative case feature with Agr-S/Tense. The resulting chain ([DP-there], t) will be a legitimated object at LF: its head, [DP-there], has all their features checked and its tail, t, gives the chain an interpretation (linked to the q-role associated with that position).

The latter -adjunction to the head- has not been taken into account in the literature. In the following paragraph, we will try to show its plausibility.

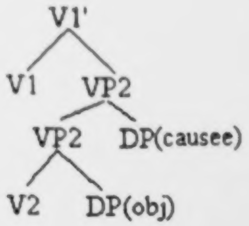
Cliticization and Case Checking. Consider a typical head-complement configuration. We have seen before that DPs have a Case to be checked with the verb. Raising to the specifier of an Agr head gives DPs the chance to indirectly enter into the checking domain of a verb and hence to check their Case features. However, a simpler option is available: adjunction of the head of DP to the verb, that is, DI. By means of DI, the Det carrying the Case feature of the whole DP will enter into the checking domain of the verb (see note 3). If this approach is correct, cliticization would be a proper means to check Case features.

The consequences such a proposal would have for the theory of Feature Checking seem very promising but so intricate that it would depart from the goals we are concerned with. As a result, I am not pursuing this matter further. In the remainder of the paper I will assume that cliticization -conceived of as an instance of DI- is a licit means of checking the Case features of DPs.

### On the Constituent Structure of Causative Constructions

The analysis of CC that will be followed here is that of Villalba (1992, 1993), which is mainly based on Li (1990) and Guasti (1992). Basically, I will assume that the constituent structure of CC is as follows:

(8)



The causative verb (V1) selects a verbal projection, VP2, as a complement. This verbal projection has an X-bar subject, the causee, which asymmetrically c-commands the internal arguments of the verbal head, V2.[5] Furthermore, V2 incorporates into V1 (plausibly) for tense reasons (see paragraph 2.3.2.). In the remainder of this section I will briefly justify such an analysis.

The Causee Is a VP-Internal Subject. This paragraph will be devoted to showing that there is enough evidence for assuming that the causee is the X-bar subject of V2, that is, the argument that asymmetrically c-commands the internal arguments of V2.

Support for this hypothesis comes from Weak Crossover (hereafter: WCO) effects. A descriptive generalization subsuming WCO effects is found in Lasnik & Stowell (1991):

(9) In a configuration where a pronoun P and a trace T are both bound by a quantifier Q, T must c-command P.

This phenomenon provides us with a clear-cut test to set up the c-commanding relations within $V2^n$. Moreover, given our hypothesis that the causee asymmetrically c-commands the internal arguments of V2, a prediction can be made: a quantifier in the causee position would share its index with a pronoun within the object position, but not conversely. The prediction is borne out, as the LF representations of the following sentences show:

(10) a. El Pere ha fet convidar tothom a la seva mare.
the Peter has made invite everybody to the his mother
'Peter has made his mother invite everybody.'
a'. *[for every x, x a person, Peter made x's mother invite x]
a". [for every x, x a person, Peter made y's mother invite x]
b. El Pere ha fet convidar la seva mare a tothom.
the Peter has made invite the his mother to everybody
'Peter has made everybody invite his mother.'
b'. [for every x, x a person, Peter made x invite x's mother]

Furthermore, wh-phrases show the same behavior:

(11) a. Qui vas fer convidar a la seva mare?
Who PAST make invite to the his mother
'Who(m) did you make his mother invite?'
a'. *[for which x, x a person, you made x's mother invite x]
a". [for which x, x a person, you made y's mother invite x]
b. A qui vas fer convidar la seva mare?
to whom PAST make invite the his mother
'Whom did you make invite his mother?'
b'. [for which x, x a person, you made x invite x's mother]

Whereas a quantifier or a wh-phrase in object position yields a WCO effect when binding a pronoun within the causee position at LF, the sentence becomes perfect in the converse configuration. It must be concluded thus that the causee asymmetrically c-commands the internal arguments of V2. There is thus enough evidence to assume our initial hypothesis: the causee is the VP-internal subject of

V2.[6]

Causative Verbs Select a Verbal Projection.

Lack of tense: There exist many empirical arguments supporting the hypothesis that causative verbs select a verbal projection as a complement. Lack of embedded tense is the most striking one.[7] Consider the following sentences:

(12) a. Ahir el Pere li va ordenar atacar el campament avui.
yesterday the Peter him-DAT PAST ordered attack the camp today
'Yesterday Peter ordered him to attack the camp today.'
b. *Ahir el Pere li va fer atacar el campament avui.
yesterday the Peter him-DAT PAST make attack the camp today
'Yesterday Peter make him attack the camp today.'

The contrast is clear-cut: CC have just one tense domain for both verbs, as a result, only one temporal adverb can be licensed.[8] If tense domains are to be derived from the presence of a Tns node, the preceding contrast might provide good evidence against analyses of CC involving an embedded Tns projection.

A similar conclusion arises if the interaction of CC and aspectual auxiliaries is taken into account:

(13) a. *El Pere va fer haver trencat el vidre al nen.
the Peter PAST make have broken the glass to-the child
*'Peter made the child have broken the glass.'
b. *La baixa temperatura va fer ser freda la sopa.
the low temperature PAST make be cold the soup
*'The low temperature made the soup be cold.'

It is not sharply clear why CC cannot contain aspectual auxiliaries. Lema (1991) suggests that aspectual auxiliaries head a SAsp(ect) projection selected by Tns. If CC have no Tns projection at all, it is impossible for SAsp to appear. Guasti (1992) arrives at a similar conclusion. It might thus be assumed that the ill-formedness of the examples in (13) is a direct consequence of the lack of an embedded Tns projection, a prerequisite for licensing an aspectual auxiliary.

Lack of agreement: It may be assumed on the basis of the evidence shown above that Catalan CC involve no Tns projection at all. However, nothing has been said about agreement. Is there an embedded Agr projection in Catalan CC? It is extremely difficult to ask this question because no trace of overt Agr appears in Catalan infinitives. Our evidence in this point cannot be direct. Nevertheless, if we assume that CC have a unique structure (this is fact the null hypothesis), the evidence from other languages sheds light on this aspect of Catalan CC. Languages having morphological causatives show no trace of an embedded Agr. Look at the following Chichewa examples from Alsina (1992):

(14) a. Nungu i-na-mú-phík-its-á máungu (kadzidzi).
9:porcupine 9:SUBJ-PAST-cook-CAUS-IND 6:pumpkins 1:owl
'The porcupine made it cook the pumpkins (the owl).'
b. Nkhandwe zi-ku-wá-mény-er-á njovu (ana).
10:foxes 10:SUBJ-PRES-2:OBJ-hit-APPL-IND 9:elephant
2:children
'The foxes are hitting the elephant for them (the children).'

Chichewa CC show the same behavior as simple transitive sentences with respect to agreement: one argument, the causee in CC and the beneficiary in benefactive applicatives, may trigger agreement with the complex verb. In general, in languages having morphological CC, the agreement properties of CC are basically the same we found in other transitive sentences. If CC involved an embedded Agr, it would be very difficult to trace any relation for this similarity. Nothing thus points to the existence of an embedded Agr in CC.

The same holds for English. Unlike a verb selecting a CP/IP (e.g. 'want' or 'believe'), the causative verb 'make' does not allow the infinitival inflection to appear:

(15) a. John makes Mary open the door.
b. *John makes Mary to open the door.
c. *John wants Mary open the door.
d. John wants Mary to open the door.

Finally, languages having inflected infinitives do not allow them to appear in CC (see Jones (1990:193) for Sardinian and Alvarez, Regueiro, & Monteagudo (1986:396) for Galician). On the basis of these facts we have evidence enough to conclude that CC have no embedded Agr projection in the above mentioned languages. However, what about Catalan? The question is to be answered theory-internally. The null hypothesis is that CC have a unique structure cross-linguistically regardless of its variation. Variation should be explained in terms of parametric differences concerning core aspects of UG. Consequently, even though we have no empirical evidence to decide whether the complement of Catalan causative verbs contains an Agr node or not, theory-internal coherence brings us to argue against the presence of such a functional projection.

Embedded Negation: Another argument against the clausal status of the causative complement is the marginality of embedded negation. There exists a sharp contrast in the following pair of sentences:

(16) a. El Pere no li va fer comprar patates.
the Peter not to-him PAST make buy potatoes
'Peter did not make him buy potatoes.'
b. ??El Pere li va fer no comprar patates.
the Peter to-him PAST make not buy potatoes
'Peter made him not buy potatoes.'

It might seem that the presence of embedded negation is, in spite of its marginality, a proof of the clausal nature of the causative complement. However, as it has been argued before (see Guasti (1992) for Italian and Villalba (1993) for Catalan and Spanish; cf. Ritter & Rosen (1993)), the negation in these constructions does not haveä a clausal status. In other words, it is an instance of constituent negation. The scope interaction of negation and quantifiers shows us that this must be so.

(17) a. El Pere desitjava no tenir molts problemes.
the Peter wanted not have many problems
'Peter wanted not to have many problems.'
a'. [for many x, x problems, Peter wanted not to have x]
a". [for not many x, x problems, Peter wanted to have x]
b. ??El Pere ha fet no venir molts alumnes.
el Pere has made not come many students
'Peter made many students not come.'
b'. [for many x, x students, Peter made x not come]
b". *[for not many x, x students, Peter make x come]

As the LF representations show, when an embedded controlled sentence is negated, either the quantified phrase -(17a')- or the negation -(17a")- may have wide scope, as it is expected for a clausal negation. However, when embedded negation under a causative verb is taken into account, only the quantifier phrase may have wide scope -(17b'). This behavior corresponds to a constituent negation status rather than to a clausal one. It may be argued then that, even though embedded negation is marginally possible under causative verbs in Catalan, it should be analyzed as an instance of constituent negation. But, why is this so? The answer is quite simple if the empirical and theory-internal evidence presented so far in this paragraph is taken into account: causative complements are verbal projections in Catalan, so they leave no chance for the presence of a clausal negation. It seems thus to be the case that, contrary to prima facie expectations, embedded negation whenever is possible argues against the clausal status of the causative complement.

To sum up, both empirical and theoretical evidence point to the hypothesis that Catalan CC select a verbal projection. Such a hypothesis will be assumed throughout.

Verb Incorporation.

Some Empirical Evidence: It has commonly been accepted in the literature that CC involve some kind of complex verb formation. Such an intuition has been set up in different ways depending on the authors and the framework. In this paper I will follow Aissen (1979), Baker (1988), Li (1990) and Guasti (1992) and assume that CC involve Verb Incorporation (hereafter: VI) of V2 to V1 in order to create a complex verb. Consider the following examples (the Swahili example is quoted from Vitale (1981)):

(18) a. mwalimu hu-wa-som-esh-a wanafunzi kurani.
teacher HAB-2:OBJ-study-CAUS-IND students Koran
b. El professor fa estudiar l'Alcorà als estudiants.
the teacher makes study the-Koran to-the students
'The teacher makes the students study the Koran.'

Looking at causativization as a Grammatical Function changing process, both languages behave exactly: the verb increases its valency with a new argument (the causer). Nevertheless, the two languages seem to have a very different syntactic correlate of this GF-changing process: whereas an affix is added to the verb in Swahili, the causativized verb is embedded under a causative verb in Catalan. However, it may be argued that the difference is less important than seems to be. Namely, on the one hand, the causative affix in Swahili is also a verb selecting a verbal projection; on the other, V1 and V2 also form a complex verb in Catalan (this is indeed the null hypothesis). Furthermore, there exists some empirical evidence pointing to such a conclusion. Consider the following sentences:

(19) La Maria ha fet acabar completament la feina al Pere.
the Mary has made finish completely the task to-the Peter
'Mary made Peter finish the task completely.'
(20) Els nens fan treballar tots la Maria.
the children make work all the Mary
'The children all make Mary work.'
(21) Els nens no faran treballar pas la Maria.
the children not make-FUT work neg the Mary
'The children will not make Mary work.'

In all three examples, V2 appears to the right of different fixed elements: a VP-adjoined adverb, a floating quantifer, and a negative particle, respectively. If, as it is commonly assumed (e.g., Belletti (1991), Pollock (1989), and Sportiche (1988)), a main verb surfacing to the right of these elements has moved from its original position, the same must be said of V2 in (19)-(21). But, where does V2 move to? If the constituent structure of CC proposed so far is assumed, there is just one (initial) landing site for the moved verb: V1. Any other move would violate the requirement that chains be minimal.[9] In fact, it seems to be the case that V2 incorporates to V1 -ex. (19)- and then they both move up to different functional heads leaving back the floating quantifier in subject position -ex. (20)- and the negative particle in the Spec (or head) of NegP -ex. (21).[10] I am not entering into details, however. For our purposes, the evidence presented so far is enough to motivate VI in Catalan CC.

Theoretical motivation and consequences: We have just seen that an analysis of Catalan CC involving VI was tenable on empirical grounds. However, something has to be said on its theoretical motivations and consequences. One theoretical motivation of VI might follow from the structure of CC assumed so far. We have seen that the causative complement had no Tns projection at all. However, as different authors have suggested (see Guéron & Hoekstra (1988), Zagona (1988)), events are to be licensed by Tns. In that case, the event argument of the embedded VP would have no chance of being licensed in situ.[11] The only way for the embedded event to be licensed is VI: once V2 incorporated, the embedded event is accessible to the main Tns and can hence be licensed. This analysis straightforwardly follows from the constituent structure of CC proposed in (8), so I will assume it to be correct. For other arguments, see Guasti (1992), Li (1990) and Villalba (1993).

What about the consequences that VI has for UG? While discussing Noun Incorporation, Baker (1988) raises the question of how complex heads behave with respect to Case. Simple heads have their Case features listed in the lexicon, but something new has to be assumed for heads formed in syntax. It seems quite plausible to assume, following Di Sciullo & Williams (1987), that Case (and other relevant) features of the simple verbs involved in VI percolate up to the complex one. Nevertheless, an immediate question arises: is feature percolation somehow restricted or is it rather a feature addition process? Baker (1988) gives overwhelming evidence against feature addition. Furthermore, he proposes the following principle constraining feature percolation:

(22) Case Frame Preservation Principle
A complex X of category A in a given language can have at most the maximal Case assigning properties allowed to a morphologically simple item of category A in that language. (Baker 1988:122)

Putting its conceptual adequacy aside, the Case Frame Preservation Principle (hereafter: CFPP) seems to be highly desirable on empirical grounds. It provides us with an explanation of the following ungrammatical sentences:[12]

(23) a. *La Maria va fer comprar un cotxe als nens al Pere.
the Mary PAST make buy a car to-the children(DAT) to-the Peter(DAT)
'Mary made Peter buy the children a car.'
b. *La Maria va fer comprar un cotxe el Pere.
the Mary PAST make buy a car(ACC) the Peter(ACC)
'Mary made Peter buy a car.'

Even though, both fer 'make' and comprar 'buy' can each legitimate an accusative and a dative DP, the resultant complex verb cannot legitimate neither two lexical datives nor two accusatives. In other words, no Case feature addition seems to be allowed at all. The CFPP thus constrains feature percolation in complex verb formation: complex verbs would bear as many case features as a simple one. This will have crucial consequences for our analysis of dative clitics (see paragraph 3.2.).

### On Clitics in Causative Constructions

Let us make a summary of the conclusions we have arrived at so far: (a) cliticization may receive a proper analysis as an instance of DI; (b) DI is a proper means to check Case features; (c) CC have the constituent structure in (8); and (d) CC involve complex verb formation -i.e., VI- under the constraints posed by the CFPP. Now we have the basic skills to return to the contrasts sketched at the beginning of this paper.

### Non-Dative Clitics.

Causee clitics: Let us consider the behavior of causees with respect to ClCl:

(24) a. *Faré dormir-los junts, els nens.
make-FUT sleep-them(ACC) together the children
b. Els faré dormir junts, els nens.
them(ACC) make-FUT sleep together the children
'The children, I will make them sleep together.'

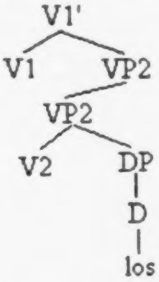
The Det head of the causee cannot incorporate to V2. As it has been shown in paragraph 2.1., the causee asymmetrically c-commands the internal arguments of V2. It must also be the case -by transitivity-that it asymmetrically c-commands V2 as well. As a result, the causee Det incorporating to V2 from the causee position would create an ill-formed chain (see note 9): the head would not c-command its trace making the derivation crash.[13] The independently motivated constituent structure of CC can easily explain why causees must climb.[14]

Internal argument clitics: transitive verbs: Consider again the following contrast:

(25) a. Faré posar-los junts a la Maria.
(I) make-FUT put-them together to the Mary
b. Els faré posar junts a la Maria.
(I) them make-FUT put together to the Mary
'I will make Mary put them together.'

Let us assume that the (simplified) constituent structure underlying both sentences is:

(26)



Now let us suppose that DI precedes VI. Det must check its Case features with an appropriate head. Given the structure in (26), three possible landing-site heads arise: V2, V1, and a higher head. However, only the former, V2, is available. The Det incorporating either to V1 or to a higher head would yield a violation of economy principles -namely, the condition that the links of a chain must be minimal (see Chomsky (1992:21)), because a more accessible landing site (i.e. V2) would be skipped. The Det must thus incorporate to V2 in order to check its Case features. Once incorporated, the Det must remain incorporated to V2, because its moving up would violate economy principles. On the one hand, incorporation to a head higher than V1 would still form a non-minimal chain, making the derivation crash. On the other hand, incorporation to V1 would now form a minimal chain. However the Det has already checked its Case features, so it has no features to check with V1. This movement would thus violate the LRP, because it is not necessary to make the derivation converge. As there is no chance for DI to apply, VI will apply incorporating the complex formed by V2 plus the Det to V1. Once there, things do not change for the Det: even if DI were still allowed to apply, Det could still not incorporate to V1, because of the LRP. Furthermore, it could not incorporate to a higher head either, because a non-minimal chain would result. The resulting complex verb will show the following structure:

(27) [V1 V1 [V2 V2+Det]]

The Det must remain adjoined to V2, a correct result for sentences like (25a).

We have just seen how the non-climbed version is the only possibility whenever DI precedes VI. Let us see what happens when it is VI that precedes DI, instead.

As it has been shown in paragraph 2.3., V2 incorporates to V1, forming a complex verb. Moreover, it was also suggested, following Di Sciullo & Williams (1987), that the features of each verb percolate up to the complex one (under the constraint of the CFPP). So then, once VI has applied, neither V1 nor V2 can further count as landing sites for the Det moving to check its Case features: the movement would not be motivated by feature checking, yielding a LRP violation. The only way for the Det to check its Case features is thus incorporation to the complex verb.[15] Note that this movement satisfies economy principles. It satisfies the LRP because it is triggered by Case checking. Furthermore, the resulting chain is minimal because feature percolation plausibly affects indices as well, so the potential intervening heads (V1, V2, and the trace of V2) do not count as heads distinct from the complex verb but rather as segments of it. The resultant complex verb would have a structure like the following:

(28) [V1 Det [V1 V1+V2]]

The Det must adjoin to the complex verb, a good result for sentences like (25b).

Let us make a summary of the analysis just suggested. We have seen how the optionality of ClCl showed in (25) is not optional at all, rather it can be derived from the interaction of two independent processes: DI -an independently motivated mechanism to check Case features- and VI -an operation at stake in CC. If DI precedes VI, the non-climbed version necessarily results -i.e. (25a). If VI precedes DI, clitics must climb -i.e. (25b).

Before closing this paragraph, let us make a brief comment on the theory-internal aspects of this approach. Leaving aside its empirical plausibility, this approach to optional ClCl seems to be plausible on theoretical grounds as well. On the one hand, because of its being the result of the interaction of two independent processes (VI and DI), no additional machinery nor stipulation is needed to account for the data. On the other hand, it gives a principled solution to the puzzle that optional ClCl posed to the LRP: optional ClCl is perfectly compatible with a principle like the LRP since no optional movement is involved at all. The apparently optional behavior of ClCl is rather the result of the order in which DI and VI apply.

Internal argument clitics: unaccusative verbs: Let us now consider the distribution of clitics heading the internal DP argument of an unaccusative verb.

(29) a. *Va fer venir-los aviat, els metges.
PAST make come-them(ACC) soon the doctors
b. Els va fer venir aviat, els metges.
them(ACC) PAST make come soon the doctors
'The doctors, he made them come soon.'

(30) a. Que va fer venir-ne molts, de metges?
that PAST make come-of.it many of doctors
b. Que en va fer venir molts, de metges?
that of.it PAST make come many of doctors
'Did he make many doctors come?'

The behavior of internal arguments of unaccusative verbs is prima facie surprising: if the Det is accusative, it must climb, whereas if partitive, it may optionally climb. From what we have seen in previous paragraphs, the expected distribution is that of (30): the Det heading an internal DP argument may climb or not depending on whether VI applies before or after DI. Namely, the sentences in (30) would receive the same analysis that those in (25). But what is this not so in (29) as well? The answer is quite straightforward for our approach to cliticization as a Case checking device. Let us consider (29a). It seems to be the case that the accusative Det has incorporated to V2 in order to check its [+acc] Case feature. However, this poses an unsolvable problem: on the one hand, venir 'to come' obviously has no [+acc] feature to be checked with; on the other, the Det cannot move up, as we have seen in paragraph 3.1.2. We have then an [+acc] feature (that of the Det) that cannot be checked, making the derivation crash. In other words, sentence (29a) cannot receive a convergent derivation. Hence, the contrast between accusative and partitive clitics with unaccusative verbs follows from simple assumptions on the structure of UG, namely, the requirement that features must be checked.

Dative Clitics. Let us finally take into account dative clitics:

(31) a. Li faré envia'ls-hi cartes.
him/her-DAT make-FUT send-them-DAT letters
a'. *'I will make them send letters to him/her.'
a". 'I will make him/her send them letters.'

The non-climbed clitic must be understood as the indirect object of V2, whereas the climbed one can only be understood as the causee. Let us see how this distribution can be accounted for under our analysis.

We have two dative DPs -the indirect object and the causee- and two heads -V1 and V2- bearing a [+dat] feature each.[16] According to this, a proper derivation for (31) would be as follows. DI precedes VI. Consequently, the head of the indirect object DP must incorporate to V2 exactly as the direct object had to (see paragraph 3.1.2. for a complete derivation). On the other hand, the head of the causee cannot incorporate to V2, as it has been shown in paragraph 3.1.1. (see also notes 11 and 15). Therefore, the causee Det incorporates to V1, where it checks its [+dat] feature. Afterward, VI applies yielding the following complex verb structure corresponding to the correct interpretation (31a"):

(32) [V1 Det(causee) [V1 V1 [V2 V2+Det(ind.obj.)]]]

To sum up, if DI precedes VI, the correct interpretation results. On the one hand, economy principles force the indirect object Det to incorporate to V2. On the other hand, the independently motivated hypothesis that the causee is the X-bar subject of V2 can straightforwardly account for the obligatoriness of causee Det climbing.

What happens with (31a')? Let us see why such an interpretation cannot be obtained. It has just been shown that the dative causee Det cannot appear on V2 regardless of its case. However, no principled reason has been suggested to prevent a Det heading a dative DP internal argument from climbing. In fact, a derivation like that of non-dative climbed Dets -i.e., VI precedes DI- prima facie seems to be possible. Nevertheless, the crucial point is that the parallelism with non-dative Dets ceases to be operative at this point. Given the analysis of VI sketched in paragraph 2.3., if VI precedes DI, the resultant complex verb will inherit the Case features of both V1 and V2 under the conditions posed by the CFPP. This had no relevant consequences for non-dative climbed Dets, because the incorporated Dets borne a different Case feature each: the CFPP allowed the complex verb to inherit enough Case features to license the non-dative internal argument and the causee -i.e. one accusative and one dative feature. However, when two dative DPs come up, the things radically change. The CFPP still allows the complex verb to inherit one accusative and one dative feature, but now these features are not enough to license both the indirect object and the causee. In other words, even though V1 and V2 bear a [+dat] feature each, the resultant complex verb will bear just one [+dat] feature, otherwise a violation of the CFPP would result. That amounts to saying that once VI applies one of the two dative Dets will not be able to check its Case feature, making the derivation crash at LF (see paragraph 2.1.). Therefore, (31a') could not receive a convergent derivation, a good result for our analysis. Moreover, such a Case-based approach to the contrast in (31) can easily be tested. It allows us to make the following prediction: if we had a dative internal argument but no causee, the former would climb without any problem. The prediction is borne out, as CC involving unaccusative embedded verbs show:

(33) a. Els hi faré arribar aquelles cartes.
them-DAT make-FUT arrive those letters
'I will mail them those letters.'
b. Aquell còctel li va fer venir gana.
that cocktail him/her-DAT make come appetite
'That cocktail whetted his appetite.'

The Det heading the dative internal argument of the unaccusative verb may climb because only one dative DP is to be licensed by the complex verb.[17] It seems thus that the existence of sentences like (33) gives us additional support for the Case-based approach to the distribution of dative clitics in Catalan CC.

## Conclusions

Our approach to the distribution of clitics in CC has rested on very simple grounds. Firstly, the constituent structure of CC gave us an answer to the obligatoriness of ClCl with external arguments. Secondly, optional ClCl was derived from the interaction of two independent processes -DI and VI- and general principles of UG. Depending on the order in which DI and VI take place, the clitic climbs or not. This optionality does not extend to internal dative clitics, since the complex verb can only license a [+dat] feature because of the CFPP. Consequently, for a CC with two dative Dets to converge DI must precede VI, following the correct distribution of dative Dets. Finally, such an analysis has proved to correctly predict the distribution of Dets heading dative internal arguments in CC with unaccusative verbs.

## NOTES

I thank the comments and help of F. Benucci, A. Branchadell, H. Hoshi, H. Lasnik, L. López, C. Picallo, and G. Rigau. Thanks also to the audiences of the CONSOLE 2 (Universitaet Tuebingen, December 1993), and the Linguistic Lunch (MIT, 24th March 1994), where parts of this paper were presented. All errors are however my own. I have also benefited from a FPI (MEC) Grant AP92.

[1] Laenzlinger (1990) and Roca (1992), following Uriagereka (1992), explicitly assume the Det to select a null *pro* as a complement. Another noteworthy option would be to consider clitics as intransitive Dets, in the way Abney (1987) treats pronouns. Anyway, the decision between one of these analysis is not relevant for the discussion. So then I am not making a choice here.

[2] According to Chomsky (1992:15) 'the category A contains B if some segment of A dominates B.

[3] It is a matter of discussion whether Case is a property of Det's or of N's (see Chomsky 1992:n. 35). I will nevertheless make the assumption -rather plausible if we look at a language like German, where the Det is inflected for Case- that it is Det rather than N that bears N-features (i.e., Case and phi-features).

[4] Indeed, a fourth logical possibility exists: adjunction to the maximal projection. Much research on this possibility, which has not been taken into account in the literature, is needed.

[5] I am assuming the standard version of c-command: A c-commands B iff A does not dominates B and every branching node dominating A also dominates B.

[6] The conclusions arrived at in this paragraph go against Li (1990), who claims that the causee and the internal arguments of V2 mutually c-command in languages like Catalan. Moreover, a corollary of this conclusion is that dative causees are not PPs but DPs. In other words, the element *a* 'to' is not a true preposition, but an inserted Case-assigner. In Branchadell (1992) it is convincingly argued that the same conclusion holds for all instances of lexical and non-lexical datives. However, for the differences between dative causees and other dative DPs see Villalba (1993).

[7] Obviously, there exist instances of CC involving a CP complement in many languages. Consider:

(i) El Pere va fer que la Maria comprés el diari.
the Peter PAST make that the Mary buy the newspaper
'Peter made Mary buy the newspaper.'

Nevertheless, sentences like that in (i) lack the bulk of characteristics making CC interesting and seem to be quite parallel to other instances of selected tensed CPs.

[8] More than one temporal adjunct can appear, indeed. However, one must be a temporal interval of the time denoted by the other:

(i) Ahir el Pere li va fer atacar el campament durant la nit.
'Yesterday, Peter made him attack the camp during the night.'

[9] For a chain CH= $[x_n, x_{n-1}, ..., x_1]$, it must be the case that, given $i > 1$, $x_i$ must c-command $x_{i-1}$, and the way from $x_{i-1}$ to $x_i$ must be minimal (shortest step); see Chomsky (1992). A 'classical' approach to these facts would rest on the ECP and relativized minimality.

[10] After V2 incorporates to V1. the latter may excorporate, at least in some

Catalan dialects (the same is true for Italian and French), allowing some adverbs, floating quantifiers and the negative particle to appear between both verbs. See Guasti (1992) and Villalba (1993).

11 It would be a matter of discussion whether CC involve two or just one event. In Guasti (1992) and Villalba (1993), it has been argued on the basis of non-temporal adverbs scope ambiguities that V2 has its own event. argument. See also Alsina (1992) for a similar conclusion; cf. with Ritter & Rosen (1993).

12 It is worth noting that some speakers find sentences like (23a) acceptable, however. See Villalba (1993) for a possible explanation.

13 It might be argued that there exists a proper representation for (24a): once VI applies, incorporating both V2 and Det to V1, the moved Det does c-command its trace. However, if, as Chomsky (1992:n.20) suggests, chain formation is better understood derivationally rather than representationally, then the violation would still be operative, ruling out the sentence.

14 However, as F. Benucci (p.c.) points out, there exist several non-standard Romance dialects where a causee clitic appears (sometimes as a copied clitic) attached to V2 (some examples can be found in Benucci 1990:ex.(15)). I have no explanation for these examples.

15 Needless to say, incorporation to a higher head is not a option, as we have just seen before. Neither is incorporation to the trace of V2. Probably, traces bear no Case features as long as it is the whole chain [V2, *t*] that percolates its features up.

16 On the nature of the case borne by the causee and the Case properties of V1 see Villalba (1993).

17 The same explanation holds for internal dative clitics in the so-called *faire-par* construction. See Kayne (1975) and Burzio (1986) for the relevant examples.

## REFERENCES

Abney, S. (1987) The English Noun Phrase in Its Sentential Aspect, PhD Dissertation, MIT.

Aissen, J. (1979) The Syntax of Causative Constructions, Garland, New York.

Alsina, A. (1992) Predicate Composition. A Theory of Syntactic Function

Alternations, PhD Dissertation, Stanford University.

Alvarez, R., X. Regueira & H. Monteagudo (1990) Gramática Galega, Editorial Galaxia, Vigo.

Baker, M. (1988) Incorporation, University of Chicago Press, Chicago.

Belletti, A. (1991) 'Generalized Verb Movement', Ms. Université de Genève.

Benucci, F. (1990) 'Ancora la destrutturazione: le perifrasi "fattitivo/percettive" in franchese e nelle lingue romanze', paper presented at the First Workshop on the Syntax of Central Romance Languages (Genève, 6-7 July 1990).

Borer, H. (1984) Parametric Syntax, Foris, Dordrecht.

Branchadell, A. (1992) Lexical and Non-Lexical Datives, PhD Dissertation, UAB.

Burzio, L. (1986) Italian Syntax, Reidel, Dordrecht.

Chomsky, N. (1992) 'A Minimalist Program for Linguistic Theory', MIT Occassional Papers in Linguistics 1, Cambridge, MA.

Chomsky, N. & H. Lasnik (1991) 'Principles and Parameters Theory', in J. Jacobs, A. von Stechow, W. Sternefield and T. Vennemann (eds.) Syntax: An International Handbook of Contemporary Research, Walter de Gruyter, Berlin.

Di Sciullo, A.M. & E. Williams (1987) On the Definition of Word, The MITPress, Cambridge, MA.

Guasti, M.T. (1992) Causative and Perception Verbs, PhD Dissertation, Université de Genève.

Guéron, J. & T. Hoekstra (1988) 'Les chaines-T et les verbes auxiliaires', Lexique, 7, 61-85

Hoffman, M. (1991) The Syntax of Argument-Structure-Changing Morphology, PhD Dissertation, MIT.

Jones, M. (1990) Sardinian Syntax, Routledge, London.

Kayne, R. (1977) Syntaxe du français, Seuil, Paris.

Laenzlinger, C. (1990) Le déplacement des clitiques et l'analyse-DP, certificat de specialisation, Université de Genève.

Lasnik, H. & T. Stowell (1991) 'Weakest Crossover', Linguistic Inquiry 22: 687-720.

Lema, J. (1991) 'Tiempo y aspecto, correlatos sintácticos y correlatos semánticos', Ms. Universidad Autònoma Metropolitana.

Li, Y. (1990) Conditions on Xo-Movement, PhD Dissertation, MIT.

Pollock, J.Y. (1989) 'Verb Movement, Universal Grammar and the Structure of IP', Linguistic Inquiry 20, 365-424.

Ritter, E. & S.T. Rosen (1993) 'Deriving Causation', Natural Language & Linguistic Theory 11, 519-555.

Roca, F. (1992) 'Object Clitics in Spanish and Catalan', Catalan Working Papers in Linguistics 1992, UAB, Bellaterra, Barcelona, 245-280.

Sportiche, D. (1988) 'A Theory of Floating Quantifiers', Linguistic Inquiry 19, 425-449.

Uriagereka, J. (1989) 'Determiner Clitic Placement', Ms. University of Maryland.

Uriagereka, J. (1992) 'Extraction Parameters', Ms. University of Maryland.

Villalba, X. (1992) 'Case, Incorporation and Economy: An Approach to Causative Constructions', Catalan Working Papers in Linguistics, UAB, Bellaterra, Barcelona.

Villalba, X. (1993) Les construccions causatives en català BA Dissertation, Universitat de Barcelona.

Vitale, A. (1981) Swahili Syntax, Foris, Dordrecht.

Zagona, K. (1988) Verb Phrase Syntax, Kluwer, Dordrecht.

Laboratori de Fonètica
Universitat de Barcelona
Gran Via de les Corts Catalanes, 585

08007 Barcelona (Spain)

ilft7@cc.uab.es

# THE DUAL STATUS OF THE NULL OBJECT IN CHINESE

**Yanfeng Qu**
**The University of British Columbia**

Abstract: This paper investigates the status of the null object in Chinese. It proposes that if the object NP is topicalized, the empty category in object position should be analyzed as a **variable**; otherwise, it is a **pro**. It argues that a pro resembles an overt pronoun in obeying Condition B, but differs from the latter in being subject to the revised version of the Generalized Control Rule initially proposed in Huang (1989). These two requirements suffice to account for the properties of the null object which Huang (1991) lists in his latest arguments against object pro. The first part of this paper is a critique of Huang's work (1991) while the second half provides additional arguments in support of the existence of an object pro.

This paper studies the null object in Mandarin Chinese. It argues that the null object in this language can be either a pro or a variable, depending on the environment. Specifically, it is a variable if an NP has moved from object position as in the case of topicalization; and it is a pro in the absence of overt NP dislocation.

## 1. Arguments Against the View that Object Pro Does Not Exist

The status of null object in Chinese has been a controversial issue in Chinese linguistics. Huang (1984, 1987, 1991) has consistently argued that there is no object pro while Xu and Langendoen (1985) and Cheng (1991) assume its existence. The central claim I am trying to establish is that there is pro in object position (cf. Hoji 1985, Saito 1992 for Japanese).[1] In this section, I will review Huang's latest arguments (1991) against object pro, and then present my counterarguments in support of its existence.

**(A). The Null Object: Only Deictic Reading**: Huang claims that the null object should be a variable, and the '**major reason**' for this postulation is the fact that it can only be interpreted deictically:

(1). $\text{John}_i$ shuo $\text{Bill}_j$ bu renshi $e_{*i/*j/k}$.
John said Bill not know
'John said that Bill does not know him/her/you/them...'

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 149-164

In contrast, an overt pronoun in the position of the null object is capable of being used deictically or anaphorically, referring to the discourse topic or to the matrix subject respectively:

(2). John$_i$ shuo Bill$_j$ bu renshi ta$_{i/*j/k}$.
John said Bill not know him
'John said that Bill does not know him.'

Huang (1991:58) argues that "if the null object were simply analyzed as a pro, its referential possibilities would be expected to be the same as those of the overt pronoun", and the interpretive contrast between the two sentences would be unexplained. He appeals to the interaction of the binding theory and a generalized version of control theory to show why pro is impossible in this position. His Generalized Control Rule (GCR) (1991:58) goes as follows:

(3) **Generalized Control Rule (GCR):**
Coindex an empty pronominal with the closest nominal element.

'Bill' in (1) is the closest nominal, but coindexing it with the null object could only result in Condition B violation. If the null object is a variable, it will be subject to Condition C (not the GCR), and it therefore cannot refer to the matrix and embedded subjects.

My response to this argument is that the fact that the null object cannot be coindexed with the matrix subject does not necessarily indicate that it must be a variable, and the solution to the disjoint reference lies right in Huang's (1991) GCR.

Let us first revise his GCR given the following example:

(4) John$_i$ gaosu Bill$_j$ [ pro$_{i/*j}$ hen xiang jianjian Mary ].
John tell Bill very like meet Mary
'John$_i$ told Bill$_j$ that he$_{i/*j}$ would like very much to meet Mary.'

Though "Bill', the indirect object, is the closest nominal, the embedded subject pro can only be co-indexed with the matrix subject 'John'. Given examples like (4), Huang's GCR should be revised as follows:

(5) **The Revised Generalized Control Rule (GCR):**
Of the c-commanding NPs, only the closest subject can be co-indexed with an empty pronominal.[2]

One of the effects of this revised GCR is that only a pro in the embedded subject position can be lexically controlled, i.e. by the subject in the next higher clause. It then follows that the disjoint reference between the **embedded** null object and the **matrix** subject in (1) above is due to the fact that the latter is not the **closest c-commanding subject**, and therefore cannot be coindexed with the object. In other words, the pro in the embedded object position is subject to two conditions: Condition B prohibits it from being coindexed with the **embedded** subject, and our revised GCR prevents it from referring to the **matrix** subject since the latter is too far away to be its licenser or controller. But there is nothing to prevent this pro from having a deictic reading as is the case of an overt pronoun. Further, the reason why an overt pronoun in the embedded object position can refer to the matrix subject is that it is not subject to GCR: what constrains it is the binding theory ONLY.

**(B). Object pro as a null epithet**: Based on a new typology of lexical categories proposed by Lasnik (1991), Huang (1991:61-61) observes that the null object in Chinese apparently has the following four properties of an epithet ([+pronominal, +referential]):

(6) a. **it cannot be A-bound**;

John$_i$ yiwei [Bill$_j$ bu xihuan e$_{*i/*j/k}$].
John think Bill not like
'John thinks that Bill does not like him/her/you...'

cf. *John$_i$ thinks that the bastard$_i$ is rich.

b. **it may be A'-bound**;

nage ren$_i$, Bill$_j$ shuo [ John$_k$ bu renshi e$_{i/*j/*k}$ ].
that man Bill said John not know
'That man, Bill said that John does not know t.'

cf. John$_i$, I saw the bastard$_i$.

c. **it need not be A'-bound**.

John$_i$ yiwei [Bill$_j$ bu xihuan e$_{*i/*j/k}$].
John think Bill not like
'John thinks that Bill does not like him/her/you...'

cf. Did you see the bastard?

d. **it may be coindexed with an argument that does not c-command the**

**epithet/the null object.**

Ruguo ni bu xihuan zhege ren$_i$, jiu buyao qing e$_i$.
if you not like that man then not invite
'If you do not like that man, then don't invite [him].

cf. The woman who met John$_i$ fell in love with this idiot$_i$.

However, despite the parallel properties listed above, there are reasons to believe that the null object does not have to be an empty epithet.

First, properties (a) and (c) are consistent with its status as a pro: it cannot be A-bound to the closest NP due to Condition B, neither to other c-commanding NPs due to the revised GCR. Further, a pro does not need an A'-binder, just like an overt pronoun with a deictic reading.

Second, suppose that the null object is a pro which is interchangeable with a pronoun unless this possibility is ruled out by independent principles (e.g. the GCR). In the sentence under property (d), the empty object can be replaced by an overt pronoun:

(7) Ruguo ni bu xihuan zhege ren$_i$, jiu buyao qing ta$_i$.
if you not like that man then not invite him
'If you do not like that man, then don't invite him.'

Here the GCR is **ineffectual** since the antecedent involved is **NOT a c-commanding NP**. In other words, the only condition governing the reference of pro in this case is Condition B, as in the case of an overt pronoun. And the easy interchangeability of these two lends strong empirical support to the pronominal nature of the null object.

Third, to say that the null object is a pro does not mean that it must be so in every case. If movement is involved, the empty category left in object position must be a variable, not a pro. And this is the case of Huang's property (b):

(8) nage ren$_i$, Bill$_j$ shuo [ John$_k$ bu renshi e$_{i/*j/*k}$ ].
that man Bill said John not know
'That man, Bill said that John does not know.'

The fact the e can be A'-bound is due to its variable status as a result of the NP dislocation. One piece of evidence comes from the subjacency effect:

(9) *nage laoshi$_k$, John$_i$ renshi [$_{np}$ [$_{ip}$ t$_i$ jian-guo e$_k$ ] de ren$_i$].
that teacher John know meet-asp DE man
'*That teacher$_k$, John knows the man [ who has met t$_k$ ].'

For comparison, there is no subjacency effect for the null object pro **embedded in an island** in the **absence** of an overtly dislocated object NP:

(10) John$_i$ renshi [$_{np}$[$_{ip}$ t$_i$ jian-guo e$_k$ ] de ren$_i$].
John know meet-asp DE man
'John knows the man who has met [him/her/it].'

Here, the lack of the subjacency effect in the last sentence provides evidence that there is no movement involved, and that the null object is a pro, not a variable.

To sum up, properties (a), (c) and (d) enumerated by Huang are consistent with our claim that the null object is a pro, and property (b) should be treated as a case of variable binding due to the movement process. Further, subjacency consideration leads us to conclude that it **must** be a pro in the absence of an overtly 'topicalized' NP.

**(C). VP-ellipsis**: The third argument in Huang (1991) is that certain cases of null objects can be analyzed as VP-ellipsis:

(11) John xihuan zheben shu, Bill ye xihuan.
John like this book Bill too like
'John likes this book, and so does Bill.'

In Huang's account, the verb 'like' in the second conjunct has been moved to the abstract INFL node, enabling the latter to L-mark (and properly govern) VP and thus allowing it to appear as an empty category. In other words, what follows the second verb is not merely a null object, but an empty VP.

This VP-ellipsis analysis may be the correct possibility for conjoined structures, but it cannot account for the null object in non-conjoined constructions. Our pro analysis is applicable for null objects in both conjoined and non-conjoined cases.

**(D). Bound pronouns**: According to the Overt Pronoun Constraint (OPC), overt pronouns cannot link to formal variables if and only if the alternation overt/empty obtains:

(12) shei$_i$ xiwang [ e$_i$/ta$_{*i/j}$ neng kanjian John ] ?
who hope he can see John
'Who hopes that he can see John?'

Based on the OPC, Huang presents his fourth argument as follows: the fact that the overt pronoun as a bound variable is prohibited in subject position is predicted since pro is allowed in this position and it wins over its overt counterpart. In contrast, only an overt pronoun is allowed in the same capacity in object position as a bound variable:

(13) shei$_i$ xiwang [ John$_k$ neng kanjian ta$_{i/j}$/(Qu: e$_{*i/j}$) ] ?
who hope John can see him
'Who hopes that John can see him?'

If there were an object pro, the overt pronoun should not be allowed by the OPC. In other words, the grammaticality of (13) with an overt pronoun just indicates that pro is not possible in this position.

Notice that this argument is based on the false premise that an overt pronoun and an empty pronoun are **interchangeable** with the same range of reference in the embedded object position. But as was discussed in section (A) above, an empty embedded object coindexed with the matrix subject is ruled out independently by the GCR since 'who' is too far away to be its possible controller, and therefore it cannot function as a bound variable (though there is nothing to prevent it from having a deictic reading). In other words, there is **no real alternation** between an empty/overt pronoun in the embedded object position since pro, due to the GCR, has fewer referential possibilities than an overt pronoun. And consequently, the OPC is irrelevant here, and the presence of an overt pronoun as a bound variable in this case is not a robust diagnostic test to preclude the presence of pro in this position.

But does the alternation obtain in the embedded subject position? The answer is yes as illustrated by (12) above: 'who' in the matrix subject position is the closest c-commanding subject, and is therefore the legitimate controller for the pro in the embedded subject position (per the GCR). Moreover, coindexing between the two does not violate Condition B either since they are in different clauses. Consequently, there is a **real alternation** between the overt and empty pronouns with the same range of referential possibilities, and only in this case then does the OPC apply: the overt pronoun gives in to its empty counterpart.[3]

**(E). Emphatic *ziji*'self':** Another argument raised by Huang against object pro is as follows:

> "it is possible to have emphatic adnominal constructions in the form of pro+*ziji*. That is, a bare *ziji* in Chinese should be analyzable either as an adnominal constructions of the form pro+*ziji*, or as a simple anaphor of the form *ziji*. Although the bare *ziji* can be used immediately after a missing subject to intensify the null subject, it cannot be used immediately after a null object to intensify the null object" (Huang 1991:70).

(14) a. (his (63)):
Zhangsan shuo [ziji hui hui jia].
Zhangsan said self can return home
'Zhangsan said that he himself can go home.'

b. (his (64)):
Zhangsan shuo [wo zhi piping ziji].
Zhangsan say I only criticize self
'Zhangsan said that I only criticize myself.'
(Not: 'Zhangsan said that I only criticized him himself, and no one else.')

Huang argues that assuming that the emphatic *ziji* is an adnominal following pro, this means that the form pro+*ziji* is possible in subject position (as in (14a)). but not in object position (as in (14b)). Therefore, there can not be object pro.

However, Huang's assumption that a bare *ziji* can be an adnominal construction is not empirically motivated. It is based on the fact that an 'overt pronoun+*ziji*' can be an adnominal construction as is shown in the following sentence:[4]

(15) (His (62))
wo zhi piping taziji.
I only criticize himself
'I only criticized him himself.'

Huang's argument is that since the 'overt pronoun+*ziji*' can be an adnominal structure as in (15), a bare *ziji* should assume the same structure as well since Chinese is a pro-drop language. Except for this argument, there is no empirical evidence in support of the ambiguous structure for a bare *ziji*.

What I propose is that *ziji* is either a simple anaphor as is commonly assumed or an adverb, and the contrast between (14a) and (14b) is precisely due to the different status of *ziji* involved. In (14a), *ziji* is used as an adverb with an embedded pro subject, for it can be located after the modal verb:

(14) a'.
Zhangsan shuo [ pro **hui ziji** hui jia].
Zhangsan said **can self** return home
'Zhangsan said that he himself can go home.'

In contrast, *ziji* in (14b) is used as a simple anaphor, not an adverb since adverbs in Chinese can never be used postverbally. Given this hypothesis, the interpretation of *ziji* in (14b) is correctly predicted by Condition A, i.e. it can only refer to the embedded subject, not the matrix one due to the intervention of the embedded subject different in person (i.e. first person) from the matrix subject (i.e. third person). Recall that Tang (1989) and Cole et al (1990) notice that an anaphor in embedded object position can refer to the matrix subject so long as the latter agrees in person with the subject of the embedded clause containing the anaphor:

(16) Zhangsan$_i$ yiwei [ John$_j$ zhi piping ziji$_{i/j}$].
Zhangsan think John only criticize self
'Zhangsan$_i$ thinks that John$_j$ only criticizes himself$_{i/j}$.'

If *ziji* here were an adnominal construction as proposed by Huang, we would have no explanation for why it can refer to both the matrix and embedded subjects since a bare object pro cannot refer to either due to Condition B and the revised GCR respectively.

**(F). *Ziji* in Idioms**: The final piece of evidence Huang uses against object pro is the distribution of the so-called generic *ziji*. The following data are his:

(17) (his (65))
Lisi shuo [ziji zuo zhi, ziji dang ].
Lisi say self do thing self be-responsible
'Lisi said that if one does a thing, then one (should) be responsible.'

(18) (his (66))
Lisi xihuan piping ziji.
Lisi like criticize self
'Lisi likes to criticize himself.'

To quote Huang (1991:70-71),

> "in (65) neither occurrence of *ziji* needs to refer to Lisi; both can be understood as having generic reference. But in (66) the postverbal *ziji* must be understood as bound by Lisi, thus lacking generic reference. Tang (1987) has argued that the so-called

> generic *ziji* is really an instance of emphatic *ziji* modifying a generic pro. On this analysis, the contrast between (65) and (66) can be reduced to the contrast between (63) and (64) [my (14a-b], namely, a subject-object asymmetry regarding the distribution of emphatic *ziji*."

Notice that the embedded clause with *ziji* in his (65) is like an idiom while the one in (66) is not. It is potentially treacherous, if not misleading, to rely on idioms to establish the distribution of generic pro since idioms tend to have strong lexical idiosyncrasies. Granted that idioms can be used, the data provided by Huang/Tang only show part of the picture: other idioms do contain *ziji* in object position:

(19) wo renwei zhe shi [ziji qipian ziji].
I think this is self cheat self
'I think that this is self-deceiving.'

Here, neither occurrence of *ziji* needs to refer to 'I'; both can be understood as generic reference, even though the second *ziji* is an **object**. If *ziji* in subject position is analyzed as an instance of emphatic *ziji* modifying a generic pro, I can see no reason why the object *ziji* cannot be analyzed likewise.

To sum up, we have examined all the five arguments Huang (1991) raised against the existence of object pro, and we reject all of them on both empirical and theoretical grounds.

## 2. More Arguments for the Existence of Object Pro

In this section, I will provide more evidence to argue that there is object pro in Chinese.

**(A). Empty Categories in Relative Clauses:** Huang (1984:545) gives the following sentence to show that the empty category in object position is a variable while the one in subject is a pro (the indexes are his):

(20) [[ $e_i$ Mai $t_j$] de $e_{*i/j}$ ] bi [[ $e_k$ zu $t_l$] de $e_{*k/l}$ ] hao.
buy DE than rent DE good
'What one buys is better than what one rents.'

In this sentence, the empty head of the HEADLESS relative clause is interpreted wrt the object, not the subject, and this, according to Huang, constitutes evidence for the variable status of the object empty category. But this sentence is not robust enough as evidence for this conclusion. Consider the following sentence:

(21) [ $e_i$ Mai $t_j$ de $e_{i/*j}$ ] bi [ $e_k$ mai $t_l$ de $e_{k/*l}$ ] jing.
buy DE than sell DE smart
'The one who sells (things) is smarter than the one who buys (things).'

Here, only the subject can be interpreted as coreferring to the empty head, presumably for semantic reasons. Following Huang's logic (1984:545), if one of the empty categories here is the variable, then the other must be a pro. This, I believe, is the case for the empty category in object position in (21): it must be an object pro since the subject is interpreted as a variable.

Considered as a whole, the relative clause data provided above, whether Huang's or mine, point to the conclusion that the chances for the object/subject empty categories to be interpreted as a variable/pro are equal so long as the resultant interpretation is acceptable.[5]

**(B). Alternation in Concessive Clauses**: There is an alternation between a pronoun and an empty category bound as a variable in sentences containing a concessive clause, as is true for the Japanese counterpart (Nishigauchi 1990:197):

(22) buguan **shei**$_i$ lai, wo dou jian **ta**$_i$/$e_i$ yixia.
no matter who come I all meet him once
'No matter who comes, I will meet with him.'

Here, the wh-phrase indirectly binds the object in the matrix clause which functions as a bound variable. What is interesting is that this object can alternate as either an overt pronoun or an empty category. Suppose that this empty category is pro. As 'who' does not c-command the matrix object, the revised GCR is inoperative. This shows that both pro and the pronoun have the same referential possibilities since both are constrained by Condition B only. Hence the licit alternation between the two. If the empty category were a variable, we would have to say that a pronoun and a variable can alternate, a highly implausible conclusion.[6]

In conclusion, the significance of this alternation is that it provides another piece of positive evidence for the proposal that the empty object MUST be a pro in certain cases.

**(C). Invalid Sources for Variables**: Consider the following sentence:

(23) John kanjian-le e.
John see-asp
'John saw e.'

If 'e' were a variable, it would have two sources: either it is base-generated, or it is derived by movement of an empty operator. But neither of these options is problem free.

**Problem One: No index at D-structure**: Assume with Saito (1985:300) that the condition that **traces (non-pronominal empty categories) must be bound** applies at D-structure. It then follows that variables cannot be base-generated. For according to Saito (1985:305), free indexing of A-positions does not take place until S-structure, and variables, if base-generated, have no index at D-structure. Therefore they cannot be bound at D-structure, and consequently, violate the condition provided above.

**Problem Two: No Subjacency Effect**: If this were a variable left by the movement of an empty operator, it should not be found in an NP complement since it is an island. But sentences with such constructions are grammatical in Chinese:

(24) John xiang wo tigong le $[_{np}$ $[_{cp}$Bill kanjian guo e] de zhengming].
John to me provide asp Bill witness asp DE evidence
'John provided me with evidence that Bill had seen [it].'

The fact that there is no subjacency violation indicates that there is no movement involved. If the null element is a pro, its grammaticality follows naturally.

**(D). Hoji's (1985:381) Argument for Japanese**: It was observed earlier that alternation exists only between pronouns and pros with the same range of referential possibilities. But there is one case where nothing prevents the two from having the same referential possibilities; curiously, a pronoun cannot be replaced by pro (cf. Hoji 1985:381 for Japanese):

(25) ?John$_i$ de mama hen xihuan ta$_i$.
John DE mother very like him
'John's mother likes him very much.'[7]

Compare with (26) below where the object is null:

(26) John$_i$ de mama hen xihuan $e_{*i}$.
John DE mother very like
'John's mother likes [him] very much.'

But as Hoji points out for Japanese, this empty element can refer to 'John' if the latter is preceded by *lian* 'even':

(27) **lian** John$_i$ de mama dou hen xihuan $ta_{?i}/e_{?i}$.
even John DE mother even very like him
'Even John's mother likes him very much.'

The fact that sentences like (27) with either the overt pronoun or pro referring to the possessor are questionable does not affect our argument here. What is relevant is that the two kinds of pronominal can alternate in the presence of *lian* 'even', thus justifying the pro status of the null object.

## 3. Ambiguity

So far, I have been claiming that an empty category in object position is analyzed as a pro in the absence of a topicalized NP in the S-initial position. This, however, does not mean to deny its variable status in some other environments. In this section, I will argue that the null object is a variable in the presence of a topicalized NP.

Notice that Huang (1984, 1987) postulates pro in subject position, but there is nothing to prevent him from treating the empty category in that position as a variable in the presence of overt subject NP dislocation.

(28) John$_i$ shuo [ $e_i$ bu renshi Bill ].
John say not know Bill
'John said that [he] did not know Bill.'

$e_i$ is a pro, subject to the GCR and Condition B.

Compare with the following sentence which contains a topicalized subject NP:

(29) Mary$_i$, John$_j$ shuo [ $e_{i/*j}$ bu renshi Bill ].
Mary John say not know Bill
'Mary, John said that [she] did not know Bill.'

$e_i$ is a variable left by NP topicalization, subject to Condition C.

Just as an empty category in subject position can be either a pro or a variable depending on the environment, there is nothing surprising for an empty category in object position to be of either of the dual capacities given the right context:

(30) John$_i$ shuo [ Bill$_j$ bu renshi $e_{*i/*j/k}$ ].
John say Bill not know
'John said that Bill did not know [him/her/you..].'

$e_i$ is a pro, subject to the GCR as well as Condition B.

Compare with the following sentence which contains a topicalized object NP:

(31) Mary$_i$, John$_j$ shuo [ Bill$_k$ bu renshi $e_{i/*j/*k}$ ].
Mary John say Bill not know
'Mary, John said that Bill did not know [her].'

$e_i$ is a variable left by NP topicalization, subject to Condition C.

My analysis further predicts that when there is one dislocated NP, the sentence with an empty category in both subject and object positions should be ambiguous with either one interpretable as a variable bound by the "topic" NP. This prediction is indeed borne out:

(32) John, wo xiang [ e yijing jian-guole e ].
John I think already meet-asp
a.'John$_i$, I think he$_i$ has met [him/her].'
b.'John$_i$, I think I or sb. else have/has met him$_i$.'

In reading (a), the subject empty category is interpreted as the variable while the object one is a pro, referring to some entity in the discourse. In reading (b), the subject empty category is a pro, controlled by the matrix subject or some discourse entity, and the object empty category has the variable status.

Look at another example. In Chinese, we can simply say:

(33) e jian-guo e.
meet-asp
'X has met Y.'

If this sentence is embedded in a relative clause, the sentence should be ambiguous given the dual nature of both empty categories. This is indeed the case:

(34) $[_{np}$ $[_{ip}$ $e_i$ jian-guo $e_j]$ de ren$_{i/j}]$ lai le.
meet-asp DE man come asp
a. 'The man$_i$ who$_i$ $t_i$ met pro$_j$ has come.'
(subject $e_i$ as the variable referring to the head noun and object $e_j$ as pro referring to some entity in the discourse)

b. 'The man$_j$ who$_j$ pro$_i$ met $t_j$ has come.'
(object $e_j$ as the variable referring to the head noun and subject $e_i$ as pro referring to some entity in the discourse)

To summarize, if there is a topicalized NP, the empty category in subject/object position should be interpreted as a **variable**; otherwise, it is a **pro** which is subject to the revised GCR as well as Condition B. Subjacency obtains in the former case if the variable is embedded in an island, but it does not with pros.

## NOTES

[1] Paul Law (1993) argues independently for the existence of object pro in Chinese, based on a different set of arguments. Zhang (1988) also argues for the existence of object pro in Chinese, but his argument is crucially based on judgement which is different from Huang's and mine.

[2] The revised GCR resembles anaphor identification proposed by Tang (1989) in that identification in both cases involves subject-orientedness.

[3] There is one inadequacy of using OPC as a test for the exclusion of pro in Chinese. In this language, **pro and reflexive** are interchangeable as a bound variable in the embedded subject position while in **the embedded object position** only **an overt pronoun and a reflexive** are interchangeable as a bound variable (cf. Huang (1991, ft 12):

(i) shei$_i$ xiwang pro$_i$/ziji$_i$/*ta$_i$ neng jiandao John?
who hope pro/self/him can meet John
'Who hopes that he can meet John?'

(ii) shei$_i$ xiwang John$_j$ neng jiandao *pro$_{i/}$/ziji$_i$/ta$_i$?
who hope John can meet pro/self/him
'Who hopes that John can meet him?'

As is shown in (i) and (ii), Chinese data are more complicated than the Romance languages where reflexives cannot alternate with pro or an overt pronoun as a bound variable. As the OPC is based on the simple two-way alternation between overt/empty pronouns in Romance languages, its application to the Chinese data where reflexives are also involved makes it less reliable as a test, and any conclusion drawn from this test should be received with caution.

[4] But notice that an adnominal construction is not the only possible structure for the 'overt pronoun+*ziji*'. It can also be a (compound) anaphor; otherwise, Condition B would be violated in the following example:

(i) ta zhi piping taziji.
he only criticize himself
'He only criticized himself.'

[5] Law (1993) also points out that headless relatives of this kind in Chinese are ambiguous.

[6] There is a mystery here: if the overt pronoun and pro alternate as a bound variable, why does the Overt Pronoun Constraint not apply, i.e. why is an overt pronoun still ok as a bound variable? I leave this question open for future research.

[7] Huang (1991:63) judges similar sentences as grammatical.

## REFERENCES

Cheng, Lisa L.-S. 1991. On the Typology of Wh-Questions. PhD Dissertation, MIT.

Cole, Peter. 1987. Null Objects in Universal Grammar, LI 18, 597-612.

Cole, Peter, Gariella Hermon and Li-May Sung. 1990. Principles and Parameters of Long-Distance Reflexives, LI 21, 1-22.

Hoji, Hajime. 1985. Logical Form Constraints and Configurational Structures in Japanese, PhD Dissertation, UW.

Huang, C-T James. 1984. On the Distribution and Reference of Empty Pronouns, LI 15, 531-574.

______ 1987. Remarks on Empty Categories in Chinese, LI 18, 321-338.

______ 1989. Pro-drop in Chinese: A Generalized Control Theory. The Null Subject Parameter, eds. by Osvaldo Jaeggli and Ken Safir, 185-214, Kluwer Academic Publishers.

______ 1991. Remarks on the Status of the Null Object. Principles and Parameters in Comparative Grammar, ed. by Robert Freidin, 56-76, MIT Press.

Lasnik, Howard. 1991. On the Necessity of Binding Conditions. Principles and Parameters in Comparative Grammar, ed. by Robert Freidin, 7-28, MIT Press.

Law, Paul. 1993. On Null Subjects and Null Arguments, Canadian Journal of Linguistics 38,1-41.

Nishigauchi, Taisuke. 1990 Quantification in the Theory of Grammar. Kluwer Academic Publishers.

Saito, Mamomu. 1985. Some Asymmetries in Japanese and Their Theoretical Implication. PhD Dissertation, MIT.

Tang, C. Jane. 1989. Chinese Reflexives, NLLT 7, 93-121.

Xu, Liejiong and D. Terence Langendoen. 1985. Topic Structures in Chinese, Language 61, 1-27.

Zhang, Shi. 1988. Argument Drop and pro. Proceedings of the Seventh WCCFL, ed. by Hagit Borer, Stanford University, CSLI 363-374.

# ON THE ORIENTATION PROBLEM IN KOREAN CAKI BINDING AND THE TYPOLOGY OF X$^0$ REFLEXIVE BINDING

Mi-Hui Cho
Indiana University

Abstract: The structural account that only a subject binds morphologically simple X$^0$ reflexives cannot explain the case where the Korean X$^0$ reflexive pronoun caki in long-distance context is bound by a nonsubject. It is shown that long-distance bound caki is actually pronominal and there are the type of verbs that restricts caki only to the pronominal use. Nonsubject binding occurs when a matrix predicate is described in terms of the object's viewpoint, namely the predicate put the Pivot on the nonsubject. Thus, it is argued that the viewpoint dimension in addition to the structural dimension should be incorporated in determining the binder for caki. Additionally, typological variation in adopting the dimensions of structure and viewpoint in X$^0$ reflexive binding is introduced.

## 1. Introduction

The Korean X$^0$ (morphologically simple) reflexive pronoun caki 'self' is bound by a local antecedent, like the English reflexive pronouns (myself, yourself, etc.) but unlike English reflexives, it is also bound by a long-distance antecedent, as shown in (1)[1].

(1) *John*$_i$*-nin* [$_{IP}$*Tom*$_j$*-i* *caki*$_{i/j}$*-lil* *chuchenha-es'-ta]-ko* *sengkakha-es'-ta.*
-NOM -NOM self-ACC recommend-PAST-DEC-COMP think-PAST-DEC
'John thought that Tom recommended self.'

The fact that caki in the embedded clause can refer back to the subject of the main clause seems to violate Binding Principle[2] A (Chomsky 1981). Principle A states that an anaphor is bound in its governing category[3], IP in this case. However, the sentence is still grammatical. Consequently, various attempts have been made to explain this non-local, 'long-distance' (henceforth LD) binding phenomenon.

Among these attempts, Yang (1984) and Wexler & Manzini (1987) propose parameterized governing categories, where the X$^0$ (morphologically simple, lexical) reflexives of Korean, Japanese, and Chinese do not have any governing categories because these languages do not have the crucial category AGR(eement). Therefore, the antecedent of caki in (1) can be local, namely the subject of the clause containing caki or LD, namely the matrix subject.

Another kind of attempt to account for LD anaphors is the X$^0$ movement analyses. These analyses originate from the assumption that anaphors undergo LF-movement to INFL (Lebeaux, 1983; Chomsky, 1986a), thus anaphors are

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 165-183

c-commanded at LF only by subjects and not by objects. This assumption has been developed by Pica (1987) and Cole, Hermon, and Sung (1990), where $X^0$ reflexives undergo successive head movement from INFL to INFL through COMP at LF, enforcing the property of obligatory subject-orientation in which LD reflexives cannot be bound by objects because they are in INFL, hence not c-commanded by anything in VP, as shown in the Chinese example in (2).

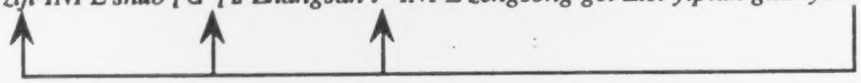
(2) from Cole & Sung (1991)

| *Wangwu$_i$* | *shuo* | *Zhangsan$_j$* | *zengsong* | *gei* | *Lisi$_k$* | *yipian* | *guanyu* | *ziji$_{i/j/*k}$* | *de* | *wenzhang.* |
|---|---|---|---|---|---|---|---|---|---|---|
| | says | | gave | to | | one | about | self | DE | article |

'Wangwu says that Zhangsan gave Lisi an article about self.'

<LF>

*Wangwu ziji-INFL shuo [$_{CP}$ [$_{IP}$Zhangsan t''-INFL zengsong gei Lisi yipian guanyu t'*



*de wenzhang.*

Since caki in (1) is assumed to move from the embedded clause to the main clause at LF, it can be bound by the LD antecedent as well as the local antecedent, as in the Chinese example.

A third account is Progovac's (1991) Relativized Subject analysis. Even though Progovac does not adopt a movement analysis, she proposes that the only subject for $X^0$ reflexives is an $X^0$ subject, namely AGR, since $X^0$ reflexives can be bound only by heads. Therefore, objects are excluded from the set of possible binders in LD binding because objects are full XP (phrasal, morphologically complex) phrases, so they cannot bind any head assuming the Structure Preserving Principle (Chomsky, 1986b). Caki in (1) can be bound by the subject of the embedded clause and the subject of the main clause because there is no AGR in Korean, thus having no governing category.

As seen above, most recent accounts try to explain the deviant behavior of $X^0$ reflexives that are bound by a LD antecedent as well as a local antecedent and they predict a striking correlation between LD binding and subject orientation such that LD anaphors exceptionlessly have the property of subject orientation. However, there are cases in which $X^0$ reflexives in LD context are bound not only by a subject but also by an object. Thus, I will claim that the **structural** account that only a subject binds $X^0$ reflexives alone cannot explain this phenomenon of nonsubject binding. Rather, I will adopt a **viewpoint** account in the pairing of the LD antecedent and $X^0$ reflexives. In section 2, I will show that the binder of caki varies depending on the type of matrix clause predicate. When a matrix predicate is described in terms of the viewpoint of the object (Object-centered predicate) in a sentence, the binder of caki is the object rather than the subject. Section 3 demonstrates that the so called LD anaphor caki is actually the pronominal use of caki evidenced by the sloppy identity test applied by Aikawa (1991) and shows that there are lexical constraints on caki where some predicates prevent caki from being bound locally. Section 4 observes the **nonsubject binding** of caki and argues that it is necessary to incorporate a viewpoint dimension in addition to a structural dimension to explain this phenomenon. Thus, I will introduce the typological variation in adopting the dimensions of structure and viewpoint in $X^0$ reflexives.

## 2. Predicate Effect

There is predicate effect in the pairing of Korean caki and an antecedent. Korean caki is bound by a subject only when the verb is malhata 'tell', as shown in (3).

(3) *John*$_i$*-i Bill*$_j$*-eke caki*$_{i/*j}$*-iy elinsicel-e tehaye malha-es'-ta.*
-NOM -DAT self-GEN childhood-about tell-PAST-DEC
'John told Bill about self's childhood.'

If we replace the verb malhata 'tell' by mutta 'ask', normally the indirect object (hereafter I will use the term "nonsubject" for antecedents other than subjects) binds caki, as shown in (4)[4].

(4) *John*$_i$*-i Bill*$_j$*-eke caki*$_{?i/j}$*-iy elinsicel-e tehaye mul-es'-ta.*
-NOM -DAT self-GEN childhood-about ask-PAST-DEC
'John asked Bill about self's childhood.'

If the subject binds caki, the sentence is odd because John is asking another person about his own life. Thus, the subject can be a possible antecedent only in abnormal circumstances, because the verb mutta 'ask' is strongly biased pragmatically toward the nonsubject as an antecedent. In the case of the verb tutta 'hear', the nonsubject also binds caki, as shown in (5).

(5) *John*$_i$*-i Bill*$_j$*-eke caki*$_{?i/j}$*-iy elinsicel-e tehaye tul-es'-ta.*
-NOM -from self-GEN childhood-about hear-PAST-DEC
'John heard from Bill about self's childhood.'

The same interpretation applied in the verb mutta 'ask' also applies in the case of the verb tutta 'hear'. On the other hand, only the subject binds caki in (6).

(6) *John*$_i$*-i Bill*$_j$*-eke caki*$_{i/*j}$*-iy chinku-il sokeha-es'-ta.*
-NOM -DAT self-GEN friend-ACC introduce-PAST-DEC
'John introduced self's friend to Bill.'

When the verb sokehata 'introduce' is replaced by the verb sokepatta 'receive the favor of introducing', the indirect object binds caki because the verb sokepatta requires the nonsubject as an antecedent for caki in (7).

(7) *John*$_i$*-i Bill*$_j$*-eke caki*$_{?i/j}$*-iy chinku-il soke-pat-es'-ta.*
-NOM -from self-GEN friend-ACC introduce-BENEF-PAST-DEC
'John received from Bill the favor of introducing self's friend.'

Thus, individual verbs can control the selection of the antecedent for caki. Therefore, the grammar must allow for both subject and nonsubject binding of caki, depending on the matrix verb.

Kuno & Kaburaki (1977) present a similar case, where each verb (predicate) selects an argument position (subject or nonsubject) whose viewpoint the speaker adopts. They call this phenomenon Empathy and the definition is given below.

(8) Empathy (Kuno & Kaburaki, 1977: 628).
Empathy is the speaker's identification, with varying degrees, with a person who participates in the event that he describes in a sentence.

The locus of empathy varies, depending on the predicate. For example, there are two kinds of verb 'give' in Japanese; yaru and kureru. The speaker must describe yaru from the subject's (giver) viewpoint (Subject-centered predicate) whereas the speaker must report kureru from the object's (recipient) viewpoint (Nonsubject-centered predicate). Thus, Subject-centered predicates give empathy focus to the subject and Nonsubject-centered predicates put empathy focus on the nonsubject. I will call empathy focus **Pivot** following Sells (1987: 455).

(9) Pivot: the one from whose point of view the report is made.

Sells describes Pivot as follows: if someone makes a report with Mary as the Pivot, that person is understood as standing in Mary's shoes. If the Pivot is located in the subject, the sentence is described in terms of the subject's viewpoint and if the Pivot is located in the nonsubject, the sentence is reported in terms of the nonsubject's viewpoint. Generally, the Pivot is located in the subject, thus the subject is the unmarked position of the Pivot following Kameyama (1984). Therefore, predicates other than those lexically specified as Nonsubject-centered predicates are Subject-centered by default.

In Korean, there are **Nonsubject-centered predicates** which give the Pivot to the nonsubject, as shown in (10).

(10) Nonsubject-centered predicates
*malhecuta* 'give the favor of telling', *mutta* 'ask', *titta* 'hear', *(toy)tollyecuta* 'return', *pillita* 'borrow', *sokepatta* 'receive the favor of introducing', *tolyeponeta* 'send back', *suyepatta* 'receive the favor of giving', ...

This inventory is not exhaustive because the Korean predicate system is productive in that the action described from the viewpoint of the referent of the subject may be converted into the action described from the viewpoint of the referent of the nonsubject by adding a benificiary morpheme, as illustrated in (11).

(11)

| Subject-centered predicates | Nonsubject-centered predicates |
|---|---|
| *malhata* 'tell' | *malhe-cu-ta* 'give the favor of telling' |
| *cuta* 'give' | *(toy)tollye-cu-ta* 'return' |
| *sokehata* 'introduce' | *soke-pat-ta* 'receive the favor of introducing' |
| *poneta* 'send' | *tollye-pone-ta* 'send back' |

In the case of causative predicates, both the object and the subject bind caki, as shown in (12) and (13).

(12) *John*$_i$*-i* *Billj-eke* *caki*$_{i/j}$*-iy* *pap-il* *mek-i-es'-ta.*
-NOM -DAT self-GEN meal-ACC eat-CAUS-PAST-DEC
'John feed Bill self's meal.'

(13) *John*$_i$*-i* *Bill*$_j$*-il* *caki*$_{i/j}$*-iy* *pang-e* *kamkim-sikhi-es'-ta.*
-NOM -ACC self-GEN room-LOC keep-CAUS-PAST-DEC
'John kept Bill in selfs room.'

Since the causative predicate is not Nonsubject-centered, it is Subject-centered by default. Thus,the subject which is Pivot binds caki. However, the object of causative predicates is the subject in deep-structure, thus the object is also Pivot. Therefore, causative predicates have two Pivots. This seems to result in the object binding as well as the subject binding of caki.

The subject which is Pivot by default becomes the best antecedent for caki, as shown in (3) and (6). However, the subject which is not Pivot because of Nonsubject-centered predicates cannot be the best antecedent and instead, the nonsubject which is Pivot is the best antecedent for caki, as shown in (4), (5), and (7). Both Pivot-hood and subjecthood participate in the determination of an antecedent. However, Pivot-hood takes precedence over subjecthood in caki binding.

(14) Pivot-antecedent principle
A Pivot binds caki.

Note that violation of the Pivot-antecedent principle does not predict a clear-cut ungrammaticality because there are the cases in which a subject non-Pivot binds caki, as shown in (4), (5), and (7), even though the meaning of the sentence is odd. Thus, we can deduce that subjecthood and Pivot-hood represent two independent dimensions: structure and viewpoint. While the structural dimension requires the subject as an antecedent, the viewpoint dimension requires the Pivot NP as an antecedent. When those two dimensions are mismatched, Pivot overrides subject in the sense of providing the preferred interpretation, but subject is not surpressed by Pivot because they are independent. Therefore, the binder for caki can be ordered as follows.

(15) Binder hierarchy for caki
Pivot and Subject > Pivot > Subject

## 3. Anaphor caki and pronominal caki

3.1 *Two-caki hypothesis:* It has been noted that morphologically simple ($X^0$) anaphors (LD anaphors) pattern with pronouns with respect to internal structure and grammatical function. Specifically, Reinhart & Reuland (1991) suggest that LD anaphors should be viewed as pronominal anaphors, thus obeying the Binding Principle B rather than A.

Aikawa (1991) proposes that there are two kinds of Japanese zibun: anaphor zibun and pronominal zibun because they behave differently with regard to the sloppy identity test. The sloppy identity test is a vehicle to test whether a pronoun or an anaphor is a bound varible. The following sentence is ambiguous.

(16) from Reinhart (1983)
*Felix hates his neighbors and so does Max.*
a. Felix hates Felix's neighbors and Max hates Max's neighbors.
b. Felix hates Felix's neighbors and Max hates Felix's neighbors.

While (16 a) is called the sloppy identity reading, thus his is a bound variable, (16 b) is called the nonsloppy (strict) identity reading. Aikawa applies this test to the Japanese reflexive pronoun zibun by adding the phrase soo-su 'do so' and predicts as follows: when zibun is locally bound, only the sloppy reading is possible, thus it is an anaphor following Williams (1977) and when zibun is nonlocally bound, both sloppy and nonsloppy readings are possible, thus it is pronominal following Reinhart (1983). This prediction is born out in (17), where zibun is bound by the local antecedent, thus the strict reading is not acceptable and zibun is an anaphor.

(17) *John$_i$-ga [Mary$_j$-ni zibun$_j$-o sono position ni suisens]-saseta. Bill$_k$-ni*
-SM -DAT self-ACC that for recommend-made -DAT
*mo soo saseta.*
too so do-made
'John$_i$ made Mary$_j$ recommend herself$_j$ for that position. (John$_i$ made) Bill$_k$ do so too.'
a. sloppy reading
John$_i$ made Mary$_j$ recommend herself$_j$ for that position.
(John$_i$ made) Bill$_k$ recommend himself$_k$ too.
b. strict reading
*John$_i$ made Mary$_j$ recommend herself$_j$ for that position.
(John$_i$ made) Bill$_k$ recommend Mary$_j$ too.

Since zibun is bound by the LD antecedent in (18), both sloppy and strict readings are possible and zibun is pronominal.

(18) *John$_i$-ga [Mary$_j$-ni zibun$_i$-o sono position ni suisens]-saseta. Bill$_k$-ni*
-SM -DAT self-ACC that for recommend-made -DAT
*mo soo saseta.*
too so do-made
'John$_i$ made Mary$_j$ recommend John$_i$ for that position. Bill do so too.'
a. sloppy reading
John$_i$ made Mary$_j$ recommend John$_i$ for that position.
Bill$_k$ made Mary$_j$ recommend Bill$_k$ too.
b. strict reading
John$_i$ made Mary$_j$ recommend John$_i$ for that position.
Bill$_k$ made Mary$_j$ recommend John$_i$ too.

The same generalization also applies in Korean. Thus, Lee (1991) adopts the two-caki hypothesis; when caki is locally bound (if it is an anaphor), only the sloppy reading is possible and when caki is nonlocally bound, both sloppy and strict readings are possible, as shown by the following examples.

(19) *John-in [ Mary$_j$-ka caki$_j$-lil chingchanha ]-tolok ha-es'-ta. Bill-to kiles'-ta.*
-NOM -NOM self-ACC praise-COMP-CAUS-PAST-DEC -too did so-DEC
'John made Mary praise self. Bill did so, too.'
a. sloppy reading
John made Mary praise Mary, and John made Bill praise Bill.
b. strict reading
*John made Mary praise Mary, and John made Bill praise Mary.

Since the local antecedent, Mary binds caki, caki is an anaphor. Thus, only the sloppy reading is possible. However, in (20), caki is pronominal because it is bound by the LD antecedent, John. Thus, both sloppy and strict readings are possible.

(20) *John$_i$-in [ Mary-ka caki$_i$-lil chingchanha ]-tolok ha-es'-ta. Bill-to kiles'-ta.*
-NOM -NOM self-ACC praise-COMP-CAUS-PAST-DEC -too did so-DEC
'John made Mary praise self. Bill did so, too.'
a. sloppy reading
John made Mary praise John and Bill made Mary praise Bill.
b. strict reading
John made Mary praise John and Bill made Mary praise John

As seen above, the sloppy identity test gives independent evidence for positing that caki really hastwo different behaviors; anaphoric and pronominal.

If we adopt the two-caki hypothesis, we do not need to explain the LD binding phenomenon as movement analyses or parameterized analyses do because the LD bound caki is not an anaphor but a pronoun, thus it obeys Binding Principle B. Note, however, that **pronominal caki** is not the same as the pronoun ki 'he' because the former must be bound by an antecedent within the same sentence, and thus is a bound pronoun whereas the latter can be bound or unbound. Only the **anaphor caki** bound by a local antecedent obeys Binding Principle A, thus the governing category need not be extended and caki need not move at LF for the pronoun caki bound by a LD antecedent. Thus, the anaphor caki and the bound pronoun caki are in complementary distribution and caki is not problematic for the Binding Principle.

However, the definition of an accessible Subject needs to be revised to determine the governing category in Korean because it is assumed that there is no AGR in Korean, thus resulting in no governing category for an $X^0$ anaphor. I will adopt Lee's (1991) rough working definition of an accessible Subject, where AGR is replaced by INFL or the Genitive marker, as shown in (21).

(21) A Subject is INFL (or Genitive marker) or the subject of an infinitive, a gerund, an NP or a small clause.

Now, we need to reanalyze the sentences containing caki according to the two-caki hypothesis. First, (1) is repeated below.

(22) *John*$_i$*-nin [*$_{IP}$ *Tom*$_j$*-i caki*$_{i/j}$*-lil chuchenha-es'-ta]-ko*
-NOM -NOM self-ACC recommend-PAST-DEC-COMP
*sengkakha-es'-ta.*
think-PAST-DEC
'John thought that Tom recommended self.'

The embedded IP is the governing category for caki because there is a governor of caki, the verb chuchenhata and the Subject, Tom. Thus, caki$_j$ bound in its governing category IP is an anaphor whereas caki$_i$ free in its governing category IP is pronominal.

The sentences (3) through (7) have the same structure, thus I repeat only (3) below.

(23) *John*$_i$*-i Bill*$_j$*-eke [*$_{NP}$ *caki*$_{i/*j}$*-iy elinsicel]-e tehaye malha-es'-ta.*
-NOM -DAT self-GEN childhood-about tell-PAST-DEC
'John told Bill about self's childhood.'

The embedded NP is the governing category for caki since there are the governor of caki, the genitive marker and the Subject, namely, the genitive marker according to (21). Because caki is free in its governing category, it is pronominal. Moreover, the fact that the pronoun ki also can occur in the position of caki , as shown in (24), confirms the claim that caki in this position is pronominal. However, the pronoun ki can be bound or unbound whereas caki must be bound in the same sentence. And ki can be bound by any NP outside the governing category regardless of Subject-hood and Pivot-hood whereas caki is bound by a subject or NP which is the Pivot.

(24) *John$_i$-i Bill$_j$-eke [$_{NP}$ ki$_{i/j/k}$-iy elinsicel]-e tehaye malha-es'-ta.*
-NOM -DAT he-GEN childhood-about tell-PAST-DEC
'John told Bill about his childhood.'

To sum up, caki bound in its governing category is an anaphor and caki free in its governing category, namely the *so* called LD anaphor is a bound pronominal. The governing category roughly corresponds to an immediate IP or NP containing caki.

3.2 *Lexical constraints on caki*: Hyams & Sigurjonsdottir (1990) claim that Icelandic sig 'self' behaves differently depending on predicates. According to them, there are lexical constraints on the $X^0$ reflexive sig. While the gefa 'give'-class verbs strongly prefer the LD antecedent, raka 'shave'-class verbs strongly prefer the local antecedent for sig, as illustrated in (25).

(25) a. *Kermit$_i$ segir ad Jon gefi$_{(subj)}$ ser$_i$ bil.*
'Kermit says that John gives SIG a car.'
b. *Jon segir ad Petur$_j$ raki$_{(subj)}$ sig$_j$*
'John says that Peter$_j$ shaves SIG.'

Therefore, they define the give-class verbs as LD verbs because sig with these verbs is strongly bound outside of the immediate clause containing sig and the shave-class verbs as local verbs because sig with these verbs is strongly bound within the immediate clause. Thus, they propose that sig with the give-class verbs is pronominal and sig with the shave-class verbs is a pure anaphor. In other words, sig is pronominal if it is bound by a nonclause-mate antecedent whereas sig is anaphoric if it is bound by a clause-mate antecedent. These two behaviors of sig are consistent with the two-caki hypothesis in Korean.

In Korean, there are also lexical constraints on caki. Caki with the give-class verbs must be bound by the LD antecedent, as shown in (26).

(26) *John$_i$-in [$_{IP}$Tom$_j$-i caki$_{i/*j}$-eke catongcha-il cu-es'-ta]-ko malha-es'-ta.*
-NOM -NOM self-DAT car-ACC give-PAST-DEC-COMP say-PAST-DEC
'John said that Tom gave self a car.'

In the normal use of the verb cuta 'give', the subject (giver) and the object (recipient) must be different. In other words, the subject gives something (DO) to somebody (IO) other than the subject. Thus, the subject in the embedded clause cannot bind caki. In this respect caki is different from sig because caki is obligatorily pronominal whereas sig is strongly preferably pronominal. Since the governing category for caki is the embedded IP, caki bound by the LD antecedent is pronominal. In the verb chotehata 'invite', the subject and the object also must have different referents, like the verb cuta. Thus, the subject in the embedded IP cannot bind caki, as shown in (27), therefore caki is pronominal.

(27) *John$_i$-in [$_{IP}$Tom$_j$-i caki$_{i/*j}$-lil choteha-es'-ta]-ko malha-es'-ta.*
-NOM -NOM self-ACC invite-PAST-DEC-COMP say-PAST-DEC
'John said that Tom invited self.'

Caki with the give-class verbs obligatorily takes a LD antecedent, thus it is pronominal. I will name the give-class verbs the **Pronominal verbs**.

There are no shave-class verbs in Korean unlike Icelandic, because inherently reflexive verbs like shave oneself and wash oneself are intransitive instead of being transitive, thus they do not have the reflexive pronoun caki as an object, as illustrated in (28).

(28) *John-i myentoha-es'-ta.*
-NOM shave-PAST-DEC
'John shaved.'

The fact that there are no shave-class verbs means that there are no verbs that constrain caki only to be anaphoric. Namely, there is no case where caki is only an anaphor, not a pronoun, in complex sentences. Thus, there are no **Anaphoric verbs** that take only a local antecedent of a complex sentence in Korean.

When the reflexive verb myentohata 'shave' is used as a transitive verb which takes caki as an object in a complex sentence, the verb becomes causative, as shown in (29).

(29) *John$_i$-nin [$_{IP}$Tom$_j$-i caki$_{i/*j}$-lil myento-sikhi-es'-ta]-ko malha-es'-ta.*
-NOM -NOM self-ACC shave-CAUS-PAST-DEC-COMP say-PAST-DEC
'John said that Tom shaved self.'

The subject of the embedded IP cannot bind caki because the verb myentosikhita 'shave someone' is causative. Causative verbs must have an object (causee), and the subject (causer) and the causee must have different referents like the give-class verbs. Therefore, the LD antecedent binds caki. Thus, the causative verb myentosikhita also belongs to the Pronominal verbs because it permits only the pronominal caki.

There are some verbs that permit both anaphoric caki and pronominal caki. In the case ofthe verb salanghata 'love', caki can be bound within the governing category, the lower IP, thus being anaphoric or it can be bound outside the governing category, thus being pronominal, as illustrated in (30).

(30) *John$_i$-nin [$_{IP}$Tom$_j$-i caki$_{i/j}$-lil salangha-n-ta]-ko malha-es'-ta.*
-NOM -NOM self-ACC love-ASP-DEC-COMP say-PAST-DEC
'John said that Tom loves self.'

Since the verb salanghata permits both anaphoric caki and pronominal caki, I will call this class of verbs **Pronominal/anaphoric verbs**. However, note that pronominal use of *caki* is much more natural with Pronominal/anaphoric verbs as in Icelandic where *sig* with the verb *elska* 'love' strongly prefers a LD antecedent.

The anaphoric use of *caki* is very weak because there are no Anaphoric verbs in complex sentences and the anaphoric use of *caki* with the Pronominal/anaphoric verbs is not prefered. Therefore, *caki* is more similar to Norwegian *seg* 'self' (Hestvik, 1989) and Danish *sig* 'self' (Vikner, 1985) than to Icelandic *sig* in that Norwegian *seg* and Danish *sig* are uniformly pronominal as claimed by Hyams & Sigurjonsdottir. Thus, Norwegian *seg* and Danish *sig*are never bound by a local antecedent but they are bound only by a LD antecedent.

In summary, I claim there are lexical constraints placed on *caki* by predicates. The Pronominal verbs permit only the pronominal use of *caki*, thus *caki* only can be bound outside the governing category, obligatorily taking a LD antecedent. The Pronominal/anaphoric verbs permit both the pronominal use and the anaphoric use of *caki*, thus *caki* can be bound outside the governing category, taking a LD antecedent and within the governing category, taking a local antecedent. There are no Anaphoric verbs which permit only the anaphoric use of *caki* in complex sentences. In (31), I give a nonexhaustive list of the verbs belonging to these two categories:

(31) a. Pronominal verbs
    all causative verbs (eg. *myentosikhta* 'shave somebody', *mekita* 'feed', *cukita* 'kill'...), *cuta* 'give', *pangmunhata* 'visit', *chepohata* 'arrest', *chotehata* 'invite', *cenhwahata* 'telephone', *annehata* 'guide', *mannata* 'meet', *ponata* 'send', *kalichita* 'teach', *pilita* 'call' ....
    b. Pronominal/anaphoric verbs
    *salanghata* 'love', *cohahata* 'like', *silhehata* 'dislike', *miwehata* 'hate', *alta* 'know', *chingchanhata* 'praise', *chuchenhata* 'recommend', *pinanhata* 'criticize', ...

Generally, while the Pronominal verbs denote an action or an activity, the Pronominal/anaphoric verbs express the static, nonactive aspect. Thus, the Pronominal verbs can support a progressive reading whereas the Pronominal/anaphoric verbs cannot. If we follow the theory of verb classification which Vendler (1967; requoted from Van Valin, 1990) originally proposes, the Pronominal verbs roughly correspond to Activities and Accomplishments because these two classes commonly involve activity and the Pronominal/anaphoric verbs correspond to States and Achievements.

Most syntactic accounts in the GB framework uniformly predict that morphologically simple ($X^0$) reflexive pronouns like Korean *caki*, Japanese *zibun*, Chinese *ziji*, Icelandic *sig*, Norwegian *seg*, and Danish *sig* are never bound by a LD antecedent which is a nonsubject but they are bound only by a LD antecedent which is a subject. Hyams & Sigurjonsdottir also prove that there is obligatory subject-orientation when Icelandic *sig* is bound by a LD antecedent, namely *sig* is pronominal. In the next section, I will show that Korean caki can be bound by a LD antecedent which is a nonsubject, unlike Icelandic sig.

## 4. Nonsubject-binding of pronominal caki

Caki is bound only by the LD antecedent which is the subject in (32) and (33), like most syntactic accounts.

(32) *John$_i$-in Bill$_j$-eke [$_{IP}$Tom$_k$-i caki$_{i/\ast j/\ast k}$-eke catongc$^h$a-il*
-NOM -DAT -NOM self-DAT car-ACC
*cu-es'-ta]-ko malha-es'-ta.*
give-PAST-DEC-COMP tell-PAST-DEC
'John told Bill that Tom gave self a car.'

(33) *Emma$_i$-ka ai$_j$-eke [$_{IP}$ ap'a$_k$-ka caki$_{i/\ast j/\ast k}$-eke nole-lil*
mother-NOM child-DAT father-NOM self-DAT song-ACC
*kalichi-es'-ta]-ko mal-ha-es'-ta.*
teach-PAST-DEC-COMP tell-PAST-DEC
'The mother told the child that the father taught self a song.'

Caki cannot be bound by the local antecedent because of the lexical constraint where the Pronominal verbs cuta 'give' and kalichita 'teach' in the embedded clause do not permit caki to be bound within the governing category, the lower IP. In the matrix clause, caki is bound only by the subject because the verb malhata 'tell' in the matrix clause is not Nonsubject-centered, thus the subject is Pivot by default and a subject which is Pivot binds caki, following the Pivot-antecedent principle in (14). In other words, the statement of the embedded clause pertains to the speaker, John and emma 'mother' in the case of the verb malhata. Thus, the speaker binds caki. However, if we change the matrix verb into an Object-centered predicate mutta 'ask' in (34) and (35), the nonsubject is Pivot, and this binds caki, unlike most syntactic accounts. This is because the statement of the embedded clause pertains to the hearer, Bill, and ai 'child' in the case of the verb mutta. Thus, the hearer binds caki.

(34) *John$_i$-in Bill$_j$-eke [$_{IP}$Tom$_k$-i caki$_{?i/j/\ast k}$-eke catongc$^h$a-il cu-es'-nya]-ko*
-NOM -DAT -NOM self-DAT car-ACC give-PAST-Q-COMP
*mulha-es'-ta.*
ask-PAST-DEC
'John asked Bill if Tom gave self a car.'

(35) *Emma$_i$-ka ai$_j$-eke [$_{IP}$ ap'a$_k$-ka caki$_{?i/j/\ast k}$-eke nole-lil kalichi-es'-nya]-ko*
mother-NOM child-DAT father-NOM self-DAT song-ACC teach-PAST-Q-COMP
*mul-es'-ta.*
ask-PAST-DEC
'The mother asked the child if the father taught self a song.'

Now, consider the case where the verb of the embedded sentence is a Pronominal/anaphoric verb in (36) and (37).

(36) *John$_i$-nin Bill$_j$-eke [$_{IP}$Tom$_k$-i caki$_{i/*j/k}$-lil miweha-n-ta]-ko malha-es'-ta.*
-NOM -DAT -NOM self-ACC hate-ASP-DEC-COMP tell-PAST-DEC
'John told Bill that Tom hates self.'

(37) *John$_i$-nin Bill$_j$-eke [$_{IP}$Tom$_k$-i caki$_{?i/j/k}$-lil miweha-n-ta]-ko til-es'-ta.*
-NOM -DAT -NOM self-ACC hate-ASP-DEC-COMP hear-PAST-DEC
'John heard from Bill that Tom hates self.'

The verb of the embedded clause miwehata 'hate' permits caki to be bound within the governing category IP because it is a Pronominal/anaphoric verb. Since the verb miwehata is not Nonsubject-centered, the subject is Pivot by default. Thus, the subject of the embedded clause with the Pivot binds caki in both (36) and (37). While the report of the embedded clause belongs to the speaker, John, with the verb malhata in (36), it belongs to the speaker, Bill, with the verb titta in (37). Specifically, the subject is Pivot by default in the case of the verb malhata in (36) whereas the nonsubject is Pivot by the Object-centered verb titta in (37). Thus, any Pivot NP binds caki.

As shown in (34), (35), and (37) the nonsubject Pivot can bind the pronominal caki, contrary to the structural accounts of most syntactic theories where the nonsubject which is a LD antecedent cannot bind an $X^0$ element like caki. This phenomenon is due to the fact that Pivot-hood which comes from the viewpoint dimension more strongly controls the pairing of an antecedent and caki than does subjecthood, which comes from the structural dimension. Therefore, we must include the viewpoint dimension in caki binding.

The addition of the viewpoint dimension in reflexive pronoun binding is not limited to Korean caki only. Japanese zibun also has this property, as shown in (38).

(38) from Kameyama (1984)
*Bill$_i$ wa John$_j$ ni [ Mary$_k$ ga zibun $_{i/j/k}$ o nikunde-i-ru koto ] o kii-ta.*
TP/SB O2 SB OB hate-PRG-PRT COMP OB hear-PST
'Bill heard from John that Mary hated self.'

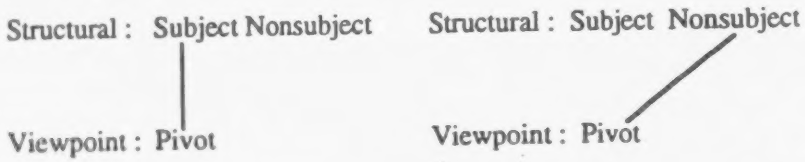
Kameyama claims that the nonsubject John can bind zibun because it has the property of Logophoricity, i.e., "the individual whose speech, thoughts, feelings, or general state of consciousness are reported or reflected in the linguistic context in which the pronoun occurs" (Clements, 1975). Namely, the sentence (38) is stated in terms of the nonsubject John's point of view. Thus, Kameyama proposes that the antecedent zibun must be a subject or logophoric individual, as shown in (39)[5].

(39) Japanese zibun : [+sub] or [+log][6]

To conclude, the structural account alone using only the concept of subjecthood is neither sufficient nor necessary to expain the binding of Korean caki and

Japanese zibun. Thus, the viewpoint dimension is introduced to account for the binding of Korean caki and Japanese zibun, as illustrated in (40).

(40) a. Subject binding  b. Nonsubject binding



Structural : Subject Nonsubject
Viewpoint : Pivot
Structural : Subject Nonsubject
Viewpoint : Pivot

When these two dimensions match, namely when the subject is Pivot, the subject is the antecedent. When they mismatch, namely when the subject is not Pivot because of the Object-centered predicates, the nonsubject Pivot is the best antecedent, since Pivothood takes precedence over subjecthood in the binding of Korean caki and Japanese zibun.

At this point, I would like to provide some examples to show how my analysis works. Three examplary sentences are given below.

(41) *John$_i$-nin Bill$_j$-eke [$_{NP}$caki$_{i/*j}$-iy chak]-il cu-es'-ta.*
-NOM -DAT self-GEN book-ACC give-PAST-DEC
'John gave Bill self's book.'

(42) *[$_{IP}$*John$_i$-nin caki$_i$-eke chak-il cu-es'-ta.]*
-NOM self-DAT book-ACC give-PAST-DEC
'John gave self a book.'

(43) [$_{IP}$*John$_i$-nin caki$_i$-lil salangha-n-ta.]*
-NOM self-ACC love-ASP-DEC
'John loves self.'

In (41), the governing category is the NP and caki is a bound pronoun because it is bound outsidethe governing category. Since the verb cuta 'give' is Subject-centered by default, only the subject John which is Pivot binds caki. In (42), the governing category is the IP and the subject John should bind caki because the verb cuta is Subject-centered. However, there is lexical constraints on caki, where caki with the Pronominal verbs like cuta cannot be bound within the governing category. Consequently, the sentence (42) is ungrammatical because caki is bound within the governing category[7]. In (43), the governing category is the IP and the subject John binds caki because the verb salanghata 'love' is Subject-centered and a Pronominal/ anaphoric verb which permits caki to be bound within the governing category.

There may be typological variation in reflexive pronouns; a language may adopt only the structural dimension or only the viewpoint dimension. Malayalam adopts only the structural dimension because possible binders must be a subject in Malayalam according to Mohanan (1982). Ewe adopts only the viewpoint dimension because possible binders must be a logophoric individual in Ewe according to Clements (1975) (requoted from Kameyama, 1984). Other languages may adopt both the structural dimension and the viewpoint dimension, like Korean caki and Japanese zibun. While Korean and Japanese allow the mismatch of these two dimensions, thus nonsubject bindings can happen, another possibility is that a language may adopt both the structural and viewpoint dimensions but does not allow the mismatch of these two dimensions, thus nonsubject bindings cannot occur. Icelandic sig belongs to the latter case according to Bresnan (requoted from Kameyama, 1984 and Sells, 1987), hence the obligatory subject-orientation in the structural account is epiphenomenal. Thus, it might be worth pursuing how these two dimensions vary in other languages like Chinese ziji[8], Nowegian seg, and Danish sig. The following chart shows the possible combinations of two dimensions and attested languages.

(44) **Typological variation of Pivot/Subject dimensions**

| Pivot | Subject | allow mismatch? | attested languages |
|---|---|---|---|
| o | o | yes | Korean, Japanese |
| o | o | no | Icelandic |
| o | _ | _ | Ewe |
| _ | o | _ | Malayalam |

## 5. Conclusion

The purpose of this paper is to demonstrate the existence of nonsubject binding of the so called LD anaphor in languages like Korean and Japanese and to give a principled account why and when it happens. Nonsubject binding occurs if the predicate of the matrix clause is described in terms of the nonsubject's viewpoint (Nonsubject-centered) in Korean. Nonsubject-centered predicates like mutta 'ask', tutta 'hear', and malhecuta 'give the favor of telling' force the statement of the embedded clause to pertain to a nonsubject. In other words, Nonsubject-centered predicates put the Pivot on the nonsubject. Therefore, caki in the embedded sentence is bound by the nonsubject, because the statement of the embedded clause is based on the nonsubject's viewpoint. Thus, from whose viewpoint the sentence is described (where the Pivot is located) is essential in caki binding. So I propose the viewpoint dimension in addition to the structural dimension. The Binder hierarchy for caki is as follows; Pivot and Subject > Pivot > Subject. The other varible in caki binding is the distinction of Proniminal and Pronominal/anaphoric verbs. While caki with Pronominal verbs is a bound pronoun which always takes a LD antecedent, caki with Pronominal/anaphoric verbs is an anaphor when it takes a local antecedent and a bound pronoun when it takes a LD antecedent. The important consequence of this paper is that the original Binding Principle (Chomsky, 1981) remains without any adaptation such as parameterized analyses or movement analyses because LD anaphors are pronominal, thus obeying Binding Principle B instead of A. Finally, the obligatory subject orientation in the Chinese example (2) seems to be due to the verbs 'give' and 'say' which are Subject-centered by default.

## ACKNOWLEDGMENT

I am indebted to Linda Schwartz for many helpful comments. Steve Franks and Natsuko Tsujimura provided helpful suggestions on the earlier version of this paper. All errors are my own.

## NOTES

[1] The special abbreviations used in this paper are as follows.
BENIF : benificiary, CAUS : causative, ASP : aspectual, Q : question

[2] Binding Principle
A. An anaphor is bound in its governing category.
B. A pronoun is free in its governing category.
C. An R-expression is free.

[3] (1) The governing category for an anaphor A is the minial category containing A, the governor for A, and the Subject accessible to A.
(2) SUBJECT : [NP, IP], [NP, NP], [AGR]
To see how this applies to (1), refer to (22) of section 3.

[4] The referent with the mark "?" means that it is semantically odd due to the mismatch of viewpoint even though it may be syntactically grammatical.

[5] Even though Sells (1987: 474) claims rather strongly that the binding of zibun is solely Pivot-oriented in Japanese, it is not so because the subject which is not Pivot is also a possible binder, as shown in (38).

[6] The term, 'logophoric individual' is equivalent to 'Pivot', here. However, I have been using the term Pivot instead of [+log] because the the term 'logophoric' has been used in Reinhart & Reuland (1991) to refer to a referent which may not be in the sentence, whereas $X^0$ reflexives like caki must be bound by an antecedent within the same sentence. The following sentence is the example of logophoric use from Reinhart & Reuland.

*The queen invited both Max and myself/me for tea.*

[7] There may be some Koreans who judge the sentence (42) to be grammatical. Caki with Pronominal verbs is never bound within the governing category in complex sentences, as shown below.

| | | | | | |
|---|---|---|---|---|---|
| *Tom$_j$-nin* | *[$_{IP}$John$_i$-i* | *caki$_{j/*i}$-eke* | *chak-il* | *cu-es'-ta]-ko* | *malha-es'-ta.* |
| -NOM | -NOM | self-DAT | book-ACC | give-PAST-COMP | say-PAST-DEC |

However, caki with Pronominal verbs may be bound within the governing category in simplex sentences like (42), even though it is not optimal but would be an alternative because caki must be bound by an antecedent within the same sentence, as discussed in section 3.1.

[8] Chinese may adopt both structural dimension and viewpoint dimension because my consultant from Taiwan shows nonsubject binding in the corresponding Chinese sentence of (35), like the Korean case. Note that Cole & Sung (1990, 1991) claim that Chinese always shows obligatory subject orientation in both local context and LD context. However, their claim is not convincing because they do not give various data that include an indirect object and always use Pronominal/anaphoric verbs in the lowest clause and Subject-centered predicates in the upper clauses, which causes obligatory subject orientation, as shown below.

*Zhangsang*$_i$ *renwei* *Lisi*$_j$ *zhidao* *Wangwu*$_k$ *xihuan* *ziji*$_{i/j/k}$.
thinks knows likes self
'Zhangsang thinks that Lisi knows that Wangwu likes himself.'

Thus, obligatory subject orientation is the result of the biased selection of verbs.

## REFERENCES

Aikawa, T. 1991. Behavior of Japanese anaphoric expressions in sloppy identity: karera 'they' and zibun 'self'. Manuscript. Ohio State University.

Chomsky, N. 1981. Lectures on government and binding. Dordrecht: Foris Publication.

———. 1986a. Knowledge of language (Ch. 3). New York: Praeger.

———. 1986b. Barriers. Cambridge: MIT Press.

Clements, G. N. 1975. The logophoric pronoun in Ewe: Its role in discourse. Journal of West African Languages 2.141-177.

Cole, P., G. Hermon & M. Sung. 1990. Principles and parameters of long-distance reflexives. Linguistic Inquiry 21.1-22.

Cole, P. & L. Sung. 1991 . Head movement and long distance reflexives. Manuscript. University of Delaware.

Hestvik, A. 1989. On the LF-movement of pronouns. Paper presented at the 6th workshop on comparative Germanic syntax, Lund University, Lund, Sweden.

Hong, K. 1987. Discourse binding of the Korean reflexive caki. Harvard Studies in Korean Linguistics 2.

Hong, S. 1985. A and A' binding in Korean and English. Doctoral dissertation. University of Connecticut, Storrs.

Hyams, N., & S. Sigurjonsdottir. 1990. The development of 'long distance anaphors': A cross-linguistics comparison with special reference to Icelandic. Language Acquisition.

Kameyama, M. 1984. Subjective/logophoric bound anaphor zibun. Chicago Linguistic Society 20.228-238.

———. 1986. Japanese zero pronominal binding: where syntax and discourse meet. Paper presented at the second SDF Japanese syntax workshop.

Kuno, S. & E. Kaburaki. 1977. Empathy & syntax. Linguistic Inquiry 8.627-672.

Lebeaux, D. 1983. A distributional difference between reciprocals and reflexives. Linguistic Inquiry 14.723-730.

Lee, Y. S. 1991. Korean caki: an anaphor and Pronoun. Manuscript. Indiana University.

Mohanan, K. P. 1982. Grammatical relations and anaphors in Malayalam. MIT working papers in linguistics 4: Papers in syntax, ed. by Marants & Stowell. MIT.

Pica, P. 1987 . On the nature of the reflexivization cycle. Proceedings of the Northeastern Linguistic Society. University of Massachusetts, Amherst.

Progovac, L. 1991. Relativized subject, long-distance reflexives, and accessibility. Manuscript. Wayne State University.

Reinhart, T. 1983. Anaphora and semantic interpretation. London: Croom Heim.

Reinhart, T. & E. Reuland. 1991. Reflexivity. Manuscript.

Sells, P. 1987. Aspects of logophoricity. Linguistic Inquiry 18.445-479.

Van Valin. 1990. Semantic parameters of split intransitivity. Language 66.221-260.

Vendler, Z. 1967. Linguistics in philosophy. Ithaca: Cornell University Press.

Vikner, S. 1985. Parameters of binder and of binding category in Danish. Working papers in Scandinavian Syntax 23.

Wexler, K. & M. Manzini. 1987. Parameters and learnability in binding theory. Parameter setting, ed. by T. Roeper and E. Williams, 41-76. Reidel: Dordrecht.

Williams, E. 1977. Discourse and logical form. Linguistic Inquiry 8.101-139.

Yang, D-W. 1984. The extended binding theory of anaphors. Theoretical Linguistic Research 1.195-218.

———. 1988. On anaphor movement. Paper presented at Northeastern Linguistic Society 19. Cornell University.

# COMPLEMENTATION OF HAUSA ASPECTUAL VERBS*

Lawàn Ɗanlādi Yàlwa
University of Kansas/UCLA/B. U. Kanô

Abstract: This paper presents a syntactic description of Hausa aspectual verbs which are traditionally called "auxiliary verbs" (cf. Jaggar 1977, Tuller 1986 and the references cited there). The paper addresses only a few facts that have not been given much attention in prior descriptions such as Jaggar 1977 and Tuller 1986. Some data that seem to show some constraints that exist between the aspectual verbs and the embedded complement is presented. It is shown that the relationship between the embedded subject and that of the matrix predicate can be analysed as involving either that of Control or Raising (cf. Perlmutter 1970, Postal 1974, and Newmeyer 1975).[1]

## 1. Introduction.

In Hausa there is a set of verbal elements that are given names in syntactic descriptions, such as auxiliary verbs (Jaggar 1977), indeterminate verbs (Parsons 1981:2 vol.2), or aspectual and modal verbs (Tuller 1986 and the references cited there). In this paper I will call them aspectual verbs (AVs). These verbal elements may appear to be raising verbs in some cases, while not in others. The subject of the complement clause of aspectual verbs in Hausa can either be arbitrary or referential, depending on the syntactic, discourse, and pragmatic factors. Aspectual verbs in Hausa express some verbal activity such as "initiation, cessation, continuation, repetition, etc. They all take VP or gerund-NP complements. A list of these verbs is given in (1) below:

(1)

| AVs. | Meanings. | AVs. | Meanings. |
|---|---|---|---|
| barì | 'stop doing/leave | cikâ | 'do too much' |
| daɗâ | 'repeat/do again' | dainâ | 'stop/cease do' |
| dingâ | 'keep on doing' | dâmu (da) | be bothered |

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 185-215

| | | | |
|---|---|---|---|
| dōsâ | keep on doing | fayê/fiyâ | do too much |
| fārâ | start/begin | fāsâ | postpone/fail |
| gamâ | finish/complete | isā | be capable |
| iyâ | be able/can | kāmâ | start/begin doing |
| ƙārâ | repeat/do again | ƙārē | finsh doing |
| kāsâ | fail/be unable to | kōmâ | return/go back |
| kumâ | repeat/do again | kusa | about to/almost |
| nēmā | try/seek to | rabu (dà) | not do as usual |
| ragē | do less as before | rasâ | be unable to do |
| riƙâ | keep on doing | sābâ (dà) | be used to doing |
| sākē | repeat/do again | sâmā | manage/succeed |
| sōmâ | start/begin | shā | do often/much |
| shigā | set about/begin/start/take on doing sth. | | |
| taɓâ | never/ever do | tārâ | share doing |
| tâsā | set about/begin/start/take steps toward doing sth. | | |
| tsayâ | persist in/stick to doing sth./persevere/be adamant | | |
| yi ta | keep doing (iteration). | | |

The syntactic function and behavior of these verbs in Hausa have received considerable attention. For instance, Parsons (1981) mentions that these verbs take dynamic, abstract, adverbial nouns, or nominal clause objects. They have general concepts of verbal modality and their lexical exponents indicate (a) single punctual actions, (b) repetitive punctual actions, or (c) durative actions. Pilszczikowa (1960) also claims that these verbs may function as auxiliary verbs taking (a) no complement, (b) noun complement, (c) verbal noun complement, or (d) a verbal theme complement. The verbal complement may be headed by a so-called "primary" verbal noun while, the nominal complement may contain a so-called "secondary" verbal noun, action noun, or other kinds of nominal expressions denoting action/activity or processes. Semantically, aspectual verbs tend to modify some verb or verbal activity/event in the complement clause-hence one of the reasons they are assumed to have no independent semantic role. The VP complement in particular may be assigned a compositional theta role (Tuller 1986). Thus, in (2):

(2) a. Aabù tā sō [$_{NP}$ wankè tukunyā]
Aabù 3fs-pst. like wash pot
'Aabù wanted to wash a pot'

b. Audu yā cê [$_{IP}$ Aabù tā tàfi --
Audù 3ms-pst. say Aabù 3fs-pst. go
[$_{NP}$ wankè tukunyā]]
wash pot
'Audù said (that) Aabu went to wash the pot'

c. Aabù tā dainà [$_{NP}$ wankè tukunyā]
Aabù she-pst. stop/cease wash pot
'Aabu stopped washing the pot'

d. [e] yà yiwu [$_{IP}$ Aabù tā ----
it 3ms-2fut. be possible/likely Aabù she-pst.
fāra [$_{NP}$ wankè tukunyā]]
start/begin wash pot
'It is possible (that) Aabù started/began washing the pot"

the underlying Activity/Event in all these sentences is Aabu's washing a pot. However, If we look at the surface syntax of Hausa aspectual verbs, it seems that they are more likely to function as Control verbs than Raising verbs. This is so because in Hausa there is no distinction between the gerund V-ing and infinitive. And the majority of these verbs do not allow lexical subjects in their complement. The few that do allow lexical subjects or null pronominal NP (Pro), only do so when the complement clause is subjunctive.[2] Furthermore, none of these verbs can take tensed complements. It is then clear that these verbs require argument subjects and may be treated as Control verbs. This fact is supported by their inability to occur in contexts like those in (3), whereas, they can occur in contexts like those in (4):

(3) a. *Audù yā dingà[$_{CP}$ (wai)[$_{IP}$ Pro/Bàlā yà tàfi]]
Audù he-pst. keep on (that) he/Bàlā he-subj. leave
'Audù kept on (that) he/Bàlā should leave'

b. *Audù yā fārà [CP Ø [IP Pro yā aurē tà]]
Audù he-pst. start/begin he-pst. marry her
'Audù he started/began he married her'

c. *Pro a-n sōmà [IP Pro à yi ruwā]
Imp.-pst. begin/start Imp-subj. do rain
'Some(one/people) started to rain' = meaning
'It started/began to rain'

(4) a. Audù yā dingà [VP gayâ ma-tà[IP Pro tà tafi]]
Audù he-pst keep on telling) to-her she-subj. leave
'Audù kept (on) telling her to leave/go'

b. Audù yā fārà [VP kar̃antà littāfi-n]
Audù he-pst. start/begin read book-ref.
'Audù started/began reading/to read the book'

c. Audù yā sōmà [NP kàr̃âtu-n littāfi-n]
Audù he-pst. start/begin reading-of book-ref.
'Audù began/started reading the book'

d. [IP Pro a-n sōmà [NP (yi-n) ruwā jiyà]]
Imp.-pst. begin/start (doing-of) rain yesterday
'It started/began raining yesterday'
'It started/began to rain yesterday'

Another issue in the description of Hausa aspectual verbs that has not received much attention is cases where a clause may contain aspectual verbs "stacked" together (to use Tuller's term). For instance, a modal verb may appear after or before an aspectual verb with a slight difference in the meaning of the EVENT/ACTIVITY contained in the complement. Consider the contrast in (5a & b) and (6a & b):

(5) a. Audù zâ-i fārà [iyà ---
Audù 1fut.-he begin/start be able
[NP (yin) kàrâtū cikin watā ɗaya]]
(doing-of) reading in month one
'Audù will start to have the ability to read in one month'

b. Audù zâ-i iya [ fārà ----
Audù 1fut-he be able begin/start
[NP (yin) kàrâtū cikin watâ ɗaya]]
(doing) reading in month one
'Audù will be able to start reading/to read in one month'

(6) a. Pro nā san zà-i fārà [ dainà ---
1s-pst. know 1fut-he begin/start stop/cease
[NP (yin) kàrâtū kâfin mū]]
(doing) reading before we
'I knew he will start to stop reading before us (=we do)'

b. Pro nā san zà-i daina [ fāra ----
I-pst. know 1fut-he stop begin
[NP (yin) kàrâtū kâfin mū]]
(doing) reading before we
'I knew he will stop to begin reading/to read before we do'

c. Kwāna-n nan Audu$_i$ zà-i fārà ---
days-of here Audù fut-he begin

[ ci gaba dà [VP faɗâ wà Balā$_j$ [IP Pro$_j$ yà fāra -
continue with telling to Bàlā he-subj. begin

[NP (yi-n) tùnâni-n [CP lōkàci-n da ----
(doing-of) thinking-of time-def. Comp.

[IP Pro$_{j/i}$ zâ-i iyà [ fārà ----
fut-he be able start
[NP gini-n gida-n-sà$_{j/i}$]]]]]]]]
building-of house-of-him
'Soon Audù will start/begin to continue telling Bàlā to start thinking about the time when he will be able to start building his house'

In (5) for instance, the (a) example indicates it is *Audù's* ability to start reading that is stressed while, in the (b) example it is the starting of the activity of the reading that is stressed. In other words, the (a) example shows certainty in the mind of the speaker while the (b) example indicates just a prediction of Audu's ability which may or may not materialize. Such stacking of modal and aspectual predicates can also be found in negative constructions as exemplified in the following sentences:

(7) a. Audù bâ zà-i iyà sākè fāra ------
Audù neg. 1fut.-he be able repeat begin/start
[NP (yi-n) kar̃ātū kâfin mū ba]
(doing-of) reading before we neg.
'Audù will not be able to repeat starting the reading before us' (i.e. before we begin the reading)

b. Audu bâ zà-i iyà sākè ƙārà -------
Audù neg. 1fut.-he be able repeat do again
[NP (yi-n) karâtū kâfin mū ba]
(doing-of) reading before we neg.
'Audù will not be able to repeat doing the reading again before us'

c. Audù bâ zà-i ƙārà sākè ------
Audù neg. 1fut.-he do again repeat
[NP (yi-n) karâtū kâfin mū ba]
(doing-of) reading before we neg.
'Audù will never again repeat the reading before us'

d. Audù bâ zà-i iyà ƙārà sākè ------
Audù neg. 1fut.-he be able do again repeat
[NP (yi-n) karâtū kâfin mū ba]
(doing-of) reading before we neg.
'Audù will never be able to repeat again the reading before us'

e. Audù bâ zà-i sākè ci-gàba dà -------
Audù neg. 1fut.-he repeat continue with
[NP (yi-n) karâtū kâfin mū ba]
(doing-of) reading before we neg.
'Audù will never again continue with the reading before us'

f. Audù bâ zâ-i ƙāra ci-gàba da ---
Audù neg.Ifut.-he do again continue with
fāra [$_{NP}$ (yi-n) kârātū kâfin mū ba]
begin/start (doing-of) reading before we neg.
'Audù will not ever continue with the initiation of the reading before us (i.e. before our act of the reading)

2. Verb/Verbal Noun Delition in the Complement Clause.

I now turn to some facts about the syntax and semantics of aspectual verbs which have not been talked about by Jaggar (1977) with the exception of the use of "Pro-form" *hakà* page 84-5 or Tuller (1986). These are cases where verbal element(s) in the complement of the aspectual verbs can be left out entirely or replaced by a resumptive verb *YI* "do". Note that, in this paper I am not concerned with the controversy of whether the verb YI is inserted or deleted at D-/S-structure (cf. Jaggar (1977:75-80) for yi-deletion and Tuller (1986, 1987) who argues against such a rule due to recoverability condition of GB Theory). Consider the examples in (8) illustrating some instances of this process:

(8) a. Mammàn Shātā yā fāra [$_{VP}$ (rērà) Bakandamiyā]
Mammàn S. he-pst. begin sing Bakandamiyā
'Mammàn Shātā began singing Bàkandamiyā a while ago'

b. Audù yā gamà [$_{VP}$ (nōmà) gōna-r̃-sa ɗâzun nàn]
Audù he-pst. finished till farm-of-him a while ago
'Audù finshed tilling his farm a while ago'

c. Audù yā kusa gamà [$_{NP}$ (nōman) gōna-r̃-sà]]
Audù he-pst about to finished tilling farm-of-him
'Audù almost finished tilling his farm by now'

d. Lādi tā dainà [$_{VP}$ (dafà) àbinci dà shà biyu-n rāna]
Lādi she-pst. stop cook food at twelve-of day
'Lādi stopped cooking lunch at twelve mid-day'

In all these sentences the verb in parentheses (and the verbal noun in (c)) may be left out and, the sense of the ACTIVITY carried out in the complement is still retained. However, not just any verbal element can be omitted. Most of the verbal elements that can be omitted involve verbs/verbal nouns or nominals that express either a phyisical ACTIVITY, ACT of SAYING, verbs that show person's OCCUPATION or someone holding a POSITION/TITTLE. Some of these activity verbals include dafâ 'cook', ɗinkâ 'sew', faɗā 'say', faskarā chop wood, fàskarē 'chopping wood, ginà 'to build, ginì building, gyārâ 'repair', karàntā 'to read', kàrâtū 'reading, ƙērâ 'smith/make', ƙīrâ 'smithing', matsayī 'status/position, mùƙāmì 'position/status', rubûtā 'to write', rùbûtū 'writing', sāƙâ 'weave', sārê 'cut off', sassàƙā 'do carpentry, shā 'drink', tārâ 'collect together, tattàrā 'collect together', etc. Note that when the verbal element is omitted, the object of the verb must be retained. The sentences in (9) illustrate more of the instances of the verbal omission:

(9) a. Mātâ-n sun gamà [$_{VP}$ (ɗinkà/sāƙa) tùmâƙasàn]
women-def. they finished sew/weave food-cover
'The women finished (sewing/weaving) the food-cover'

b. Aabù tā fārà [$_{NP}$ (fàɗa-r̃) lâbāri-n]
Aabù she-pst. start/begin telling-of story-ref.
'Aabù started (telling) the story'

c. Pro nā ci gàba dà [$_{VP}$ (karàntà) Ruwan Bagajā yâu]
1pl-pst. continue with read Ruwan Bagajā today
'I continued (reading/to read) Ruwan Bagajā today

d. Shēhù$_i$ yā dainà [$_{NP}$ (sàrautà-r̃) ----
Shēhù he-pst. cease/stop chieftainship-of
Gàlâdīmâ à Kanō]
Gàlâdīmâ in Kanō
'Shēhù ceased (being/ruling) as the Gàlâdīmâ in Kanō'

In these sentences, it is natural to omit the elements in the parentheses since the objects are the prototypical objects of the omitted elements (cf. Dixon 1991). Similarly, this omission is only possible in cases where the information or activity expressed in the complement is known by the addressee. If the addressee cannot determine the type of action or is unable to infer the information about the agent of the action, such omission is not possible. Furthermore, the activity must be known (by the interlocutors) to be the type which the matrix agent is known to do on a regular basis. In other words, the matrix agent must be the same as that of the complement. With verbs like *sò* 'like/want/love, *sayar̃ (dà)* 'sell', *mântā* 'forget, *mallakā* 'possess', etc., such verbal omission is impossible, since these verbs are virtually unrestricted in the types of complements they can take. To see this consider these examples:

(10) a. Lādi tā sō [VP/NP *(ginà/gini-n) gida-n-ta]
Lādi she-pst. like build/building-of house-of-her
'Lādi liked (to build/building) her house'

b. Aabù tā mântā dà [NP *(ɗība-n/kwāshè) ---
Aabù she-pst forgot with dipping out-of/removing
bùlâ-n ...]
blocks-ref
'Aabù forgot (dipping out/removing) the blocks ...

c. Sōjōji-n sun ƙi [VP *( tsallàkà) ---
soldiers-def. they-pst. refuse/hate jumping over
gadà-r̃]
bridge-ref.
'The soldiers refused/hated (to jump/jumping over) the bridge'

These examples illustrate the fact that we cannot omit the verb in the complement if the matrix agent is not known by the addressee to do such kind of activity. However, we can omit the verbs in the complements of sentences in (11) due to prior knowledge of the kind of inherent activity done by the matrix agent:[3]

(11) a. Bir̃kilàn$_i$ yā fārà [NP/VP (ɗībà-n/kwāshê)--
brick-layer he-pst. begin taking-of/removing
bùlà-n]
blocks-ref.
'The brick-layer started (taking/removing) the
blocks'

b. Shātā yā fārà [VP/NP (rērà)/(wāƙa-r̃) ---
Shātā he-pst. begin/start sing/song-of
Bàkandamiyā]
Bàkandamiyā]
'Shātā began/started (singing) the Bakandamiyā
(song)'

c. Mâtā-t-ā tā gamà[VP (dafà/gir̃ki-n)--
wife-of-me she-pst. finish cooking/cooking-of
àbinci-n rāna]
food-of-day
'My wife finished (cooking) lunch'
#My wife finished *eating* lunch'

Note that when the verb in (11c) is ommitted, the sentence does not mean "my wife finished eating the food". To get this meaning the verb should be followed by the associative preposition da to form the complex verb *gamâ dà* "finshed with" (see next paragraph).

Another feature of aspectual verbs that is not discussed by Tuller (1986) and is mentioned only in passing in a footnote by Jaggar (1977) is their ability to take a complement introduced by the preposition *dà* with'. The meaning of the activity/event in this complement is different from the one without the preposition. Not all of the aspectual verbs listed in (1) above can take this associative preposition. Consider the examples in (12):

(12) a. Lādi tā fārâ [pp dà [vp (dafà) àbinci---
Lādi she began with (cooking-of) food
kâfin/bāyan ....]]
before/after
'Lādi began with the cooking of the food
after/before ......'

b. Ɗâlibâ-ni sun fārâ [pp dà [vp (karànta) --
students-def. they-pst. began with (read)
Ruwan Bagajā kâfin .....]]
Ruwan Bagajā before
'The students began with the reading of Ruwan
Bagajā before..'

The complements in these examples describe an event/activity which the agent of the action has done or is going to do before or after undertaking another activity. That is why the complement is usually followed by a temporal clause. Some of the aspectual verbs that take this kind of complement are the following:

(13)

| AVs. | Meanings. | AVs. | Meanings. |
|---|---|---|---|
| daɗâ (dà) | 'repeat/do again' | dâmu (dà) | 'be bothered' |
| fārâ (dà) | start/begin | gamâ (dà) | finish/complete |
| ƙārâ (dà) | repeat/do again | ƙārè (dà) | finsh doing |
| kōmâ (dà) | return/go back | kusa (dà) | about to/almost |
| ràbu (dà) | not do as usual | sābâ (dà) | be used to doing |
| sōmâ (dà) | start/begin doing | | |
| tāshi (dà) | set about/begin/start/take steps toward doing sth. | | |
| tsayâ (dà) | persist in/stick to doing sth./persevere/be adamant | | |

Some aspectual verbs can also be used with a grade change to give a change in meaning. Note that there are seven grades of verbs in Huasa and such a grade system is used in to classify verbs into morphologically distinct forms on the basis of of final termination either in vowel or vowel followed by a consonant and tonal pattern. The termination of the verb determines the kind of object it may take. Most aspectual verbs are of grade I or IV, but a few appear in grade II. Two such verbs are *fārâ* and *sōmâ* Their use is illustrated in the following:

(14) a. Rumfa yā fâri $[_{NP}$ gini-n gānuwā à Kanō]
Rumfa he-pst. start building-of city-wall in Kanō
'Rùmfa initiated the building of Kanō wall'

b. $[_{CP}$Wâ$_i$ $[_{IP}$ t$_i$ ya sōmi $[_{NP}$ kāmà gwagwârē ---
who he began catch unmarried
à gàri-n nan]]]?
at town-of this
'Who initiated the catching of gwagwârē (unmarried people) in this town'?

Another issue in this section involves instances where the aspectual verbs occur in intransitive constructions. Here, the surface agent in the matrix clause expresses/describes an ACTIVITY, EVENT or STATE. Consider the examples in (15):

(15) a. Wâsā yā tāshi $[_{CP}$ bāyan $[_{IP}$ faɗâ yā ɓarkē --
game it-pst. end after fight it-pst. bust
à bikî-n]]
at ceremony-ref.
'The games/play ended after the fight broke out at the ceremony'

b. kishi-n Aabù yā fārâ $[_{CP}$ lōkàci-n dà -----
jealousy-of Aabù 3ms. began time-def. Comp.
$[_{IP}$ Pro ta ga miji-n-tà târe dà Rābi]]
she-pst. see husband-of-her together with Rābi
'Aabù's jealousy began when she saw her husband with Rābi'

It will be interesting to find out if these sentences involve any syntactic raising of the subject of the complement clause to the matrix subject position. I turn to this issue now.

## 3. Possibilities for Raising to Matrix Subject Position.

As I mentioned earlier, there are very few modal auxiliary verbs that take complement sentences. The sentences in (16) are therefore impossible where the senential subject is either at its D-structure position or extraposed:

(16) a. *[$_{CP}$ cêwā [$_{IP}$ Audù yā gamà aikì]] yā fārâ
that Audù he-pst. finish work it-pst. begin
'That Audù finished the work has began'

b. *[$_{CP}$ cêwā [$_{IP}$ Aabù zā ta àuri Ali]] yā ---
that Audù fut-she marry Ali it-pst.
kusa
almost/near
'That Aabù will marry Ali is near/almost'

c. *Audù$_i$ yā ci-gaba da [$_{IP}$ t$_i$ yà zama wāwā]
Audù he-pst. continue with he-subj. become idiot
'Audu continued to be an idiot'

d. *[e] yā fārâ [$_{CP}$ cêwā [$_{IP}$ Audù yā gamà aikì]]
3ms-pst. begin/start that Audù he-pst. finish work
'It began/started that Audù has finished work'

e. *[e] yā kusa [$_{CP}$ cêwā [$_{IP}$ Aabù zā tà àuri Ali]]
3ms-pst. almost/near that Audù fut-she marry Ali
'It is almost near (i.e. the time) that Aabù will marry Ali'

An existential NP cannot occur as subject of these verbs as the following examples show:

(17) a. *Akwai fārà [$_{NP}$ rûɗāmi-n siyāsâ à Najēriyā]
there's begin confusion-of politics in Nigeria
'There is an initiation of political confusion in Nigeria'

b. *Akwai daina [$_{NP}$ sāmù-n dāma-r̃ nasarâ ---
there's cease getting-of chance-of victory
à wâsà-n]
in games-ref.
'There cease to be a chance for victory in the games'

Another class of sentences that seem not to allow raising of the embedded complement is the impersonal passive (a construction where the the person acting as the subject of the sentence is not expressed). The examples in (18) are all impossible sentences in Hausa (note the final vowel of **amincē** which indicates the form it takes with direct object noun in (18a) and also the passive forms of the verbs meaning 'be believed', of the sentence in (18b)):

(18) a. *Pro sun amincè Audù$_i$ [$_{IP}$ t$_i$ ya zama sar̃kī]
they-pst. believe Audù he-subj. become emir
'They believed Audù to be the Emir'

b. *Audù$_i$ yā àmintu/yàr̃du [$_{CP}$ (dà) -------
Audù he-pst. be believe ??
[$_{IP}$ t$_i$ yà zama sar̃kī]]
he-subj. be/become emir
'Audù was/is believed to be the Emir

As we can see from these examples impersonal passive cannot be expressed as in (18b). The only way to render a impersonal passive construction in Hausa is to use an impersonal subject with the subject of the complement in its D-structure position or realised as a dative object of the matrix predicate as illustrated in (19):

(19) a. Pro a-n amincē [$_{CP}$ (cêwā) ----
Imp-pst. believe/agree/permit that
[$_{IP}$ Audù yà zama sar̃kī]]
Audù he-subj. be/become emir
'It is/was believed/permitted that Audù become Emir'

b. Pro an amincē **wà** Audù$_i$ ------
Imp-pst. believe **dat**. Audù

$[_{CP}$ (dà/cēwā) $[_{IP}$ t$_i$ yà zama sar̃kī]
that he-subj. be/become emir
'Audù is/was believed/permitted that he be Emir'
'Audù is/was permitted/allowed to become the Emir'

The presence of a complementizer in (19b) clearly shows that Audù is not the subject of the complement. Nor is it the derived object of the matrix predicate. This dative object can be focussed or questioned as in (20):

(20) a. $[_{CP}$ Audù$_i$ (nē) $[_{IP}$ Pro$_j$ a-kà amincē wà t$_i$ --
Audù cop. Imp-rel.-pst. believe dat.
$[_{CP}$ (dà) $[_{IP}$ Pro$_i$ yà zama sar̃kī]]]]
that he-subj. be/become emir
'It is/was Audù who was/is permitted to be the Emir'

b. $[_{CP}$ wâ$_i$ $[_{IP}$ Pro$_j$ a-kà amincē wà t$_i$ ---
who Imp-rel-pst. believe dat.
$[_{CP}$ (dà) $[_{IP}$ Pro$_i$ yà zama sar̃kī]]]]
that he-subj. be/become emir
'Who was/is believed/permitted to become the Emir'?

If we insist on raising analysis, then a sentence like (21) below would have either of the deep structures in (b-c) both with null INFL in the complement clauses. And to arrive at the surface structure, we need to do at least two operations. If the structure in (b) is the deep structure, then we have to introduce a rule that inserts the appropriate subject agreement clitic. If however structure (c) is assumed, all we need to do is to move the subject Audù to the matrix subject position deriving (22):

(21) a. Audù yā fārà $[_{NP}$ (yi-n) kūkā]
Audù he-pst. begin/start (doing-of) crying
'Audù began crying'

b. [ [e] fārà [$_{I''}$ Audù [$_{I'}$ [$_{I}$ Ø] [$_{VP/NP}$ yi-n kūkā]]]]
c. [ [e] yā fārà [$_{I''}$ Audù [$_{I'}$ [$_{I}$ Ø] [$_{VP/NP}$ yi-n kūkā]]]]

(22) Audù$_i$ yā fārà [$_{I''}$ t$_i$ [$_{I'}$ [$_{I}$ Ø [$_{VP/NP}$ yi-n kūkā]]]]
Audù he-pst. begin doing-of crying
'Audù began crying'

But as Tuller (1986:489) correctly claims there is no problem with this derivation, since the object in the embedded VP/NP can gets its Case from the genitive linker **-n**, meaning 'of', which links two nouns (either simple nouns or verbal nouns) to form one complex nominal phrase. But where there is no Case assigner the object NP cannot get Case violating Case filter. Even if the verb **fārà** is able to assign Case, such Case will percolate only to the head of IP. And it will remain there since the INFL lacks any lexical element (AGR or TNS) to transmit the Case down to the object.

Furthermore, one might argue that the subject complements of the aspectual verbs like those in (15) above are raised from their D-structure positions (indicated with a trace of the raised NP) in either (23) or (24) below due to the inability of the aspectual verbs to assign a theta role to these NPs. The problem with this analysis is the fact that the subject positions of some of these sentences may be theta positions. If this is so, then this movement will be a violation of the theta criterion.[4] In fact the meanings of the sentences in (15) is different from those in (23-24) which are actually ungrammatical with the intended meanings of those sentences in (15):

(23) a. [e] a-n **tāshi** (yi-n) wâsâ-n [$_{CP}$ bāyan -
Imp.-pst. set about doing-of game-ref. after
[$_{IP}$ faɗâ yā ɓarkê à bikî-n]]]]
fight it-pst. burst/break at ceremony-ref.
'They/someone (prepared to) get involved in playing the game after the fight has started/broken out at the ceremony'

b. *Pro$_i$* su$_i$-n tāshi wâsâ-n [$_{CP}$ bāyan
they-pst. set about doing games-def. after
[$_{IP}$ faɗâ yā ɓarkê à bikī-n]]]]
fight it-pst. burst/brake at ceremony-ref.
'They set about to play the games after the fighting broke out at the ceremony'

c. *mutânē$_i$* su$_i$-n tāshi (yi-n) wâsâ-n -----
people they-pst. get into doing-of games-def.
[$_{CP}$ bāyan[$_{IP}$ faɗâ yā ɓarkê à bikī-n]]]]
after fight it-pst burst/brake at ceremony-ref
'Some people/they are about to begin the games just after the fighting has broken out at the ceremony'

d. mutânē su-n tāshi [$_{PP}$ dàgà wâsâ-n ---
people they-pst. end doing from games-ref.
[$_{CP}$ bāyan [$_{IP}$ faɗâ yā ɓarkê à bikī-n]]]]
after fight it-pst burst/brake at ceremony-ref
'Some people/they have left from (the activity of) the games after fighting has broken out at the ceremony'

(23)' a. Pro a-n tâshi wâsâ-n[$_{CP}$ bāyan
Imp-pst. set about doing/end games-ref. after
[$_{IP}$ faɗâ yā ɓarkê à bikī-n]]]]
fight it-pst. burst/brake at ceremony-ref.
'Some people/they got into/ended the games after the fighting has broken out at the ceremony'
#The games was stopped after a fight has broken out at the ceremony'

b. mutânē su-n tâshi wâsâ-n
people they-pst. set about doing/end games-ref.
[$_{CP}$ bāyan [$_{IP}$ faɗâ yā ɓarkê à bikī-n]]]]
after fight it-pst burst/brake at ceremony-ref
'People got into/ended (the activity of) the games after the fighting has broken out at the ceremony'

c. mutânē sun **tâshi** [$_{NP}$ (yi-n) wâsà-n ---
people they-pst. set about doing-of games-ref.
[$_{CP}$ bāyan [$_{IP}$ faɗâ yā ɓarkē à biki-n]]]]
after fight it-pst burst/brake at ceremony-ref
'People have just decided to get into the
games/play after the fighting has broken out at the
ceremony'

d. ?**wâsâ-n**$_i$ yā **tâsu** t$_i$ [$_{CP}$ bāyan ---
games it-pst. get into doing after
[$_{IP}$ faɗâ yā ɓarkê à biki-n]]]]
fight it-pst. burst/brake out at ceremony-ref.
'The games ended/was stopped after the fighting
has broken out at the ceremony'

(24) a. Pro yā fārâ kīshi-n Aabù -----
3ms-pst. begin jealousy-of Aabù
[$_{CP}$ lōkàcī-n dà[$_{IP}$ Pro ta ga miji-n-ta ---
time-def. Comp. 3fs-rel-pst. see husband-of-her
târe dà Rābi]]
together with Rābi
'Aabù's jealousy began when she saw her husband
with Rābi'

b. ?kīshi-n Aabù$_i$ yā **ci-gàba** t$_i$ ---
jealousy-of Aabù it-pst. continue
[$_{CP}$ lōkàcī-n dà [$_{IP}$ Pro ta ga miji-n-tà ------
time-def. Comp. 3fs-rel-pst see husband-of-her
târe dà Rābi]]
together with Rābi
'Aabù's jealousy continued when she saw her
husband with Rābi'

c. *kīshi-n Aabù$_i$ yā **dàinu** t$_i$ [$_{CP}$ lōkàcî-n dà
jealousy-of Aabù it-pst. cease/stop time-def. Comp
[$_{IP}$ Pro ta ga miji-n-tà târe dà Rābi]]
3fs-rel-pst see husband-of-her together with Rābi
'Aabù's jealousy ceased/stopped when she saw her
husband with Rābi'

It is for this reason that I am assuming that this position could be regarded as an argument position and also a theta-position. Note that there is no ambiguity in interpretation of (23a). The sentence in (23a) exemplifies an active impersonal sentence in which "a" + (TNS) appears to function as an identifier of a subject with arbitrary reference, roughly equivalent in interpretation to arbitrary PRO or 'one' of English or 'on' of French and similar pronouns in other Romance languages (see Jaeggli 1986; Rizzi 1986, and Hyams 1983, 1986). I would like to propose that this "a" , the impersonal marker, functions in the same way as the definite pronominal clitics and that the feature [+/- specifed] should be included (in Hausa) as one of the identifiers of Pro (see Tuller 1982,1984, 1987, and Yalwa 1992). To conclude this section, we have so far observed that Hausa aspectual verbs can take different types of complements (as discussed above), including some complements that indicate the state of a situation or event.[5] I will now turn to the issue of Control in relation to the position of the subject of the complement clause of these verbs.

## 4. The Null Subject of the Complement of Aspectual Verbs.

With regard to the issue of Control of the null argument of the complement of infinitive and gerund clauses, Manzini (1983) argues that big PRO is a pure anaphor. She outlines numerous important properties of PRO that a Control theory must account for. She proposes that when PRO happens to occur in a sentential complement, it must be coreferential with the subject or with the object of the matrix clause. Bouchard (1984) on the other hand argues that PRO is either a pronoun or an anaphor depending on the context. That is, in essence, Control theory is derivable from the binding theory (also cf. Chomsky 1981, 1982, 1986; Borer 1983, 1988; Lebeaux 1984). I don't intend to summarize the issues raised in these works, but rather to show that the subject of the NP/VP complement of Hausa aspectual verbs may be the empty category, big PRO while, that of the aspectual verbs taking IP complement is the so called small Pro.[6]

Let me first begin with the discussion of the non-aspectual verbs taking IP complements. The first thing to note is that in Hausa, a small Pro can also be controlled. In the examples in (25) below, we can see that the subject position of the embedded clause is occupied by Pro which is coindexed with its identifier(=Agr in INFL). This Pro is controlled by the subject *Audu* in (25a), by the object *Lādi* in (25b). It is then clear that the [NP, IP] of the embedded clause is controlled by an argument in the matrix clause. This is possible due to the fact that the Pro is fully identified by Agr, which happens to refer to an argument higher in the structure. If the Agr in the embedded clause does not have the same set of features of identification with that of the matrix clause, Control is impossible from either the subject or the object of the higher clause, except where there is an IMPLICIT ARGUMENT (see Rizzi 1986, Hyams 1986, Brody and Manzini 1986, and Yalwa 1992) in the matrix clause as exemplified in (26b):

(25) a. Audù$_i$ yā ƙōƙàrtā[CP Ø [IP Pro$_i$ ya tàfi]]
Audù he-pst. try he-subj. leave/go
'Audu tried to leave/go'

b. Audù$_i$ yā shā-kân Lādi$_j$ [CP Ø ----
Audù he-pst. persuade Lādi
[IP Pro$_j$ tà tafi]]
she-subj. leave/go
'Audu persuaded/convinced Ladi to leave/go'

(26) a. Audù$_i$ yā ƙōƙàrtā (IA)$_j$ [CP Ø ----
Audù he-pst. try (they)
[IP Pro$_j$ sù tàfi gidā]]
they-subj. go home
'Audu tried for (them) to go home'

b. Audù$_i$ yā ƙōƙàrtā **ma-sù$_j$** [CP 0 ---
Audù he-pst. try **dat-them**
[IP Pro sù tàfi gidā]]
they-subj. go home
'Audu tried for them to go home'

In these examples, Pro is governed by AGR which makes it impossible to [illegible] a big PRO. The argument that controls the Pro depends very much on the lexical requirements of the verb. In other words, the C- and/or S- selection of the verb concerned. Therefore, I assume that in these cases. INFL does not have to raise obligatorily. Similarly, I assume that the Pro in these structures is actually a pronominal and it is subject to the Binding Principle B (see Yalwa 1992 for extensive discussion on this issue).

I now return to the main concern of this paper. That is the Control of the null subject of the complements of Hausa aspectual verbs. First let me briefly give examples of the aspectual verbs I mentioned right from the begining of the paper that take sentential complements. Some of these aspectual verbs include the verbs: (a) **isā** " suffice/be capable/be able"; (b) **iyâ** " be able/can"; (c) **rigā** already/have done"; (d) **barì** "let"; (e) **sābâ (dà)** "used to"; (f) **dàmu (dà)** " worry about/disturb"; (g) **ràbu dà** "not do for some time"; and verbs of saying and liking such as (h) **cê** "say"; (i) **sô** 'want/like/love"; (j) **gayâ** "tell" etc. These verbs prefer to have the subject of the embedded clause to have a coreferential interpretation with an NP in the matrix clause. Consider the sentences in (27-28) as illustrations of the obligatory Control found with these predicates:

(27) a. Audu$_i$ yā barì Bàlā$_j$ [IP Pro$_j$ yà tàfi gidā]
Audù he-pst let/allow Bàlā (he)-subj. go home
'Audu let/allowed Balā to go home'

b. Gwamnâ$_i$ yā isa [IP Pro$_i$ yà yankè hukunci]
governor he-pst be capable (he)-sub. cut verdict
'The governor(suffices)/is capable to give a verdict

c. Mālàmi yā sâ ɗālibā$_j$ [IP Pro$_j$ tà yi aikì]
teacher he-pst make student she-subj. do work
'The teacher made the (female) student do the work'

d. Sāni$_i$ zâ-i iyâ [IP Pro$_i$ yà yi aiki-n yau]
Sani 1fut-he be able he-subj. do work-ref. today
'Sani will be able/can do the work today'

(28) *a. Audu$_i$ yā bar̃ Bàlā$_j$ [IP Pro$_i$ yà tàfi gidā]
Audù he-pst. let Bàlā he-subj. go home
'Audu let/allowed Bàlā to go home"

*b. Gwamnâ$_i$ yā isa ----
governor he-pst. be capable/suffice
[IP Bàlā$_j$ yà yankè hukuncì]
Bàlā he-sub. cut verdict
'The governor(suffice)/is capable for Bàlā to give a verdict'

*c. Sāni$_i$ zâ-i iyâ [IP Bàlā$_j$ yà yi aiki-n]
Sani 1fut-he be able Bàlā he-subj. do work-ref.
'Sani will be able/capable for Bàlā to do the work

It seems to me that the obligatory coreference observed in these sentences has something to do with the semantics/ theta grids of the verbs as indicated earlier. But when we turn to the Hausa aspectual verbs and also some motion verbs taking VP/gerund complements, we will see that the Control we find with verbs like *ƙōƙàrtā* "try" is also observed here.[7] In Hausa, the raising analysis is really not easy to explain, since there is no infinitive in Hausa. Aspectual verbs as well as motion and Control verbs in Hausa may take nominalized complements or sentential complements depending on the type of the predicate used (see Yalwa forthcoming). Intuitively, one can infer that the matrix subject is the same subject that has performed the ACTIVITY/EVENT contained in the embedded complement. For instance, in (29) it seems that evidence for Control is much higher than that of raising:

(29) a. Audù$_i$ yā sō [$_{NP}$ PRO$_i$ tàfiyâ gidā]
Audù he-past want leaving/going home
'Audu wanted to go home/going home'

b. Dirēbà-n$_i$ yā daina [$_{NP}$ PRO$_i$ tūƙà mōtà-r̃]
driver-def. he-pst. stop/cease driving car-ref.
'The driver stopped driving the car'

c. Shâgâri$_i$ yā cigàbā dà[$_{NP}$ PRO$_i$ zāmā shūgàba-n-mu]
Shâgâri he-pst. continue with becoming leader-of-us
'Shâgâri continued to be our leader'

In these examples the matrix subject is the same as the embedded one. What we need is to coindex them at S-structure and nothing more can be added. Note that the matrix subject position can take different types of NPs. The subjects can be generic, specific, plural or sentential. The subject can also be null. The examples in (30) indicate that lexical NPs can occur as subjects of aspectual verbs:

(30) a. Bàlā$_i$ yā fārà [$_{NP}$ PRO$_i$ màllakà-r̃ mōtōcī/mōtā]
Bàlā he-pst. begin owning-of cars/car
'Bàlā began to own/have cars/a car'

b. Littàttàfa-n Lārabcī$_i$ sun dainà ----
books-of Arabic they-pst. cease/stop
[$_{NP}$ PRO$_i$ sâmuwā à Los Angeles]
being available in Los Angeles
'Arabic books ceased to be available/found in L.A.'
'Arabic books can no longer be found in L.A.'

c. Wâsa-n dambe$_i$ yā cigàba dà ---
game-of boxing it-pst. continue with
[$_{NP}$ PRO$_i$ sāmù-n fari-n jinī à Kanô]
getting-of popularity in Kanô
'Boxing continued to be gaining popularity in Kanô

d. Fàrfēsâ$_i$ yā cigàba dà [$_{NP}$ PRO$_i$ sāmù-n farin jinī]
professor he-pst. continue with getting-of popularity
'The professor continued to be popular/liked' or
'The professor continued getting popularity'

e. Rìga-r̃ ruwā$_i$ tā dainà [$_{NP}$ PRO$_i$ sàyuwā à Kanō]
gown-of rain it-pst. cease/stop being sold in Kanō
'Rain coat stopped to be/ceased being sellable in
Kanò'

These examples seem to indicate that any lexical NP whether it is concrete, specific, generic, or plural/collective NPs can occur as subjects of Hausa aspectual verbs with no constraint on what might occur in the complement.[8]

## 5. Conclusion.

In this paper I tried to describe and provide an analysis on the complementation pattern of Hausa aspectual verbs. I have provided some instances of aspectual verb complementation that has not been given attention or mentioned in passing in some of the previous works on the syntax of Hausa aspectual verbs. I also tried to show that syntactically, the phenomenon of Control in this type of complementation exist in Hausa. That is the so-called small Pro and big PRO in Hausa can be Controlled. And I argued that big PRO is not available in IPs with morphological INFLs (i.e. whether it is +tense + Agr= all other TAMs, or -tense +Agr = subjunctive). The [NP, IP] position can be governed. The big PRO is found only in clauses lacking lexical INFL such as the complements of aspectual verbs and the NP/VP complements of controlled predicates. It is hoped that this work will stimulate other people working on Hausa syntax and semantics to do more research on this poorly understood group of Hausa verbs.

## NOTES

* Part of this paper was presented at the 23rd ACAL held at Michigan State University, East Lancing in 1992. Another version of it was also

presented at the UCLA department of Linguistics Colloquium in May 1993, where most of the material in this paper was presented and discussed. I thank the participants for the stimulating discussion that led to this work. I would like to particulaly thank Professor R G. Schuh (the Chair of my dissertation committee) for the valuable comments on the issues discussed in this paper. This paper would not have been possisble without his encouragement and support. All remaining errors and omissions in the paper are the responsibility of the author

[1] In TGG as well as in GB, it is assumed that aspectual and modal verbs do not assign theta role to the subject position of their complement clauses. Some linguists like Perlmutter (1970) argue that aspectual verbs like '*begin* may be transitive or intransitive. The transitive one being a control verb while, the intransitive one a raising verb. But this dual analysis is rejected by Newmeyer (1975) who considers only raising analysis for English auxiliary and modal verbs. I am not concern with the controversy of whether the structures are control or raising. I am interested in providing the facts only. Of course, in some cases I will discuss such issues should there be any need for that

[2] According to Parsons 1981:261, 'a subjunctive can be used in place of a verbal noun ..... at least the accusative ones'. He cites the following examples.

(i) a Pro$_i$ a-n sâ su$_j$ [Pro$_j$ sù yi nōmā]
Imp.-pst. make them 3pl.-subj. do farming
'They were made to farm/to do farming'

b Pro$_i$ yā fārà [Pro$_i$ yà yi zāgī]
3ms-pst. begin/start 3ms-subj. do insult
He began to insult/abuse'

c. Pro$_i$ yā kōyā wà ɗa-n-sà [Pro$_i$ yà yi sātā]
3ms-pst. teach dat. son-of-him 3ms-subj. do theft
'He taught his son how to still'

d. Pro$_i$ sun ƙi [Pro$_i$ sù yi aikì]
3pl.-pst. refuse/fail 3pl-subj. do work
'They refused to work'

While I find the sentences in (a and c-d) to be grammatical, I find (b) somewhat unusual. In my dialect, such an aspectual verb can only govern a VP or NP complement with this TAM (cf. Newman and Schuh 1974 and Schuh 1985 for discussion of this term). One possible instance where one can say (b) (according to one consultant) that is where a child for instance is just beginning to learn how to insult others. There are however some aspectual verbs that can take subjunctive complements. These include **sābâ** 'used to', **nèmā** (in the sense of *try*', **isā** be capable of', **dâmū** be bothered'. and **ràbū** 'not do for some time'. For examples (cf. Jaggar 1977:82). With **iyâ** 'be able', which Jaggar included in his examples, I find the sentence acceptable only if the TAM of the matrix clause is the 'future' especially second future, or habitual and, much better if the subjunctive clause is followed by a kind of adverbial clause as in:

(ii) a. Pro$_i$ yâ iyâ [Pro$_i$ yà rina ƙyallâ-n]
he-2fut. be able he-subj. dye cloth-ref
'He can/may be able to dye the cloth

b. *Pro$_i$ yā iyâ [Pro$_i$ ya rina ƙyallâ-n]
he-pst. be able he-subj. dye cloth-ref.
'He was able to dye the cloth'

c. ?Pro$_i$ yakàn iyâ [Pro$_i$ yà rina ƙyallàn [idan lōkaci bai ƙurê ba
'He is/able (habitually) to dye the cloth when it is not too late

These examples illustrate that we need to study more about the syntax of the relationships between the TAMs and aspectual verbs. Some of the aspectual verbs differ with others w.r.t. the semantics and type of complements they take. I leave this issue here.

[3] Note further that in (10), if the verbs are omitted, the expected meanings will change. That is it will not mean "making/performing rather, it will mean liking, forgeting or hating what is described in the NP-object complement. In (11) however, the omitable verbs are all expected in this slot. And there will be no change in meaning when they are omitted. It seems that omission is possible when the the action in the complement is a kind of a description of the person (i.e. agent indicating his/her day to day habitual action or where the previous context is known by the addressee. This is why the VP as a whole can be omitted as in the following examples:

a. Mai tuwô-n tā fārâ [$_{VP}$ (V NP) dà wuri]
owner-of tuwō-def. she began/started with early
The tuwō-cook or preparer/maker began/started early

b. Shātā yā gamâ [$_{VP}$ (V NP) dab dà [$_{NP}$ zuwà-n 'yan rawân]]
Shātā he finished just about/near coming-of dancers
Shātā finished just before the coming of the dancers

c. Mālàmi-n yā gamâ [$_{VP}$ (V NP) kâfin [$_{NP}$ zuwà-n ɗâlibâ-n]]
teacher-def. he finished before coming-of students-ref
'The teacher finished before the coming of the students'

[4] If, as Tuller (1986) argues, the subject positions of these predicates are [$_{NP}$ e ], in that there is nothing occupying these positions, and they are non-Theta positions, one would expect an argument somewhere in the embedded clause to freely move into that position. Similarly, one may also try to resort to a kind of ad hoc explanation by stipulating that

when an NP is moved into the subject position, then there is an obligatory rule that will delete the indefinite subject, '*an*', and replace it with an appropriate agreement marker. But this could not explain why we cannot derive (b) from (a):

a. [$_{IP}$ [e] a-n dôki Bintà]
Imp.-Pst. hit Bintà
'Someone hit Binta' or 'Binta was hit'

b. *Bintà$_i$ tā/a-n dōki t$_i$
*'Binta (she)/(one) hit (her)'

[5] There are however some problems when the TAM of the main clause is continuative or when the complement of the aspectual verb is stative. The following examples show that both continuative and perfective TAMs can be used in the matrix clause whether the complement is a VP or NP (no matter what the verbal predicate is in the complement). It must be mentioned that the sentences with the perfective TAM are much better than those with the continuative.

(i) ?a Audù yanâ fārà [$_{PRO}$ sāmùn gidā]
Audù he-is beginning/starting to get a house

?b Audù yanâ fārà [$_{PRO}$ màllakà-r̃ gidā]
'Audù he-is beginning/starting to own a house'

c. Audu yanâ fārà [$_{PRO}$ ginà/ginin gidā]
Audu he-is beginning/starting to build/building a house'

Here, we can see that ginà 'to build' and ginì 'building' express ACTION (a situation involving a change), whereas sāmû 'getting/having' and màllakâ 'having/possessing' express STATE (a situation with no change). Predicates with stative expressions whose matrix TAM is progressive are much better if followed by certain adverbial expressions as in the following:

(ii) a. Manajōjin kamfànîn$_i$ sunâ fārà [PRO$_i$ sāmùn [$_{NP}$ gidâjen
sù$_i$/kânsù$_i$ [$_{AP}$ (bāyan shêkarâ biyar)]]]
'The managers of the company start/begin getting
(owning) their houses after five years'

b. Manajōjin kamfànîn$_i$ sun fārà [$_{NP_i}$ sāmùn [$_{NP}$ gidâjen
sù$_i$/kânsù$_i$ [$_{AP}$ (bāyan shêkarâ biyar)]]].
'The managers of the company started/began getting
(owning) their houses after five years'

If the embedded predicate is a stative verb, the sentence becomes ungrammatical. Consider the following:

(iii) *a Balā yā daina [PRO kwance/tsàye/zaune]
Balā he-pst. cease/stop being lying/standing/sitting

?*b Balā yā kāsa [ PRO kwance/tsàye/zàune]
'Balā he-is unable (be in a state of) lying/standing/sitting

This is one of the instances where aspectual verbs differ from the continuative aspect marker which may take a stative verb/noun as in the following example (cf. Tuller 1986:466):

(iv) Balā yanā [PRO [VP V[e] kwance/tsàye/zàune]].
'Balā he-is (in a state of) lying/standing/sitting

Note that if the statives are conjoined (in certain stylistic expressions), the sentence becomes grammatical as in the following examples:

(v) a. Balā duk yā kāsa [PRO zaune] kuma
Balā adv he-pst be unable be seated and
[Pro yā kāsa [PRO tsàye]]
he-pst be unable be standing
'Balā was unable to sit and was unable to stand (statively)
Balā did'nt know what to do'

b. Balā yā kāsà [PRO zàune dà tsàye].
Balā was unable to sit and stand up' i.e.
Balā did'nt know what to do

[6] I must mention that I assume the (extended) projection principle as in Chomsky (1981 and 1986). I also agree with Chomsky's (1986) proposal that there is an agreement between the SPEC and head where he assumes that they share an abstract 0-feature F, between the specifier of XP and its head. This relation is formulated as in the following:

SPEC - head Agreement:

In [ X '' ... YP ... [X ' ... X ... ]].
Where X agrees with YP.

[7] Note that in English, the differences in semantic selection between raising and control structures in infinitival embedded clauses is best exemplified in the following sentences:

a. [Audù has gone [COMP[PRO to see the manager]]]
b. [Audù is going [Ø[t to see the manager]]]

In (a) the control structure, the PRO is coindexed with the D-struture subject Audu at S-structure. In (b) however, the D-structure subject of the embedded clause is raised/moved to the subject position of the matrix clause after the S-bar deletion leaving a trace (cf. Lamiroy 1987).

[8] The examples below indicate that this is not so, in that certain predicates in the VP are incompatible with certain types of matrix subject NPs but still I assume that this has nothing to do with the referential properties of the PRO:

*a. Wāsan ƙwallō yā fāra [PRO sāmu-n dāɗi à L.A.]
'Soccer it-pst begin/start getting-of enjoyment in L.A.
'Soccer began to enjoy (= be accepted by people) in L.A.'

b. Rigar̃ ruwā tā daina [PRO ?*ɓàtuwā/ɓatà a Kanō]
'Rain coat it stopped/ceased to disappear in Kanō'

It is not clear to me whether the problem here has something to do with the relationship between the matrix subjects and the predicates in the VP/NP or, the problem has to do with the semantics of the aspectual verbs and the mentioned subject, in the sense that the former constraint the types of predicates that can occur as its argument.

## REFERENCES

Borer, H. (1983) Parametric Syntax. Foris Publications.

Borer, H. (1988) "Anaphoric AGR", UC Irvine MS.

Bouchard, D. (1984) On the Content of Empty Categories. Foris, Dordretch.

Brody, M. and Manzini, R. (1986) "On Implicit Arguments", MS.

Chomsky, N. (1981) Lectures on Government and Binding. Foris, Dordretch.

Chomsky, N. (1982b) Concepts and Consequences of the Theory of Government andBinding. MIT Press, Cambridge, Mass.

Chomsky, N. (1986) Knowledge of Language: Its Nature, Origins and Use. Praeger, N.Y, London.

Dixon, R. M. W. (1991) A New Approach to English Grammar on Semantic Principles .....Clarendon Press. Oxford.

Hyams, N. M. (1983) The Acquisition of Parameterized Grammar. CUNY, Ph.D dissertation.

Hyams, N. M. (1986) "The Interpretation and Acquisition of Italian Impersonal Si". In Syntax and Semantics, Volume 19: The Syntax of Pronominal Clitics", edited by Borer, H.

Jaeggli, O. (1986) "Case, Doubbled NPs and Extraction", Syntax and Semantics 19.

Jaggar, P. (1977) "The Nature and Function of Auxiliary Verbs in Hausa", In Papers in Chadic Linguistics, edited by P. Newman and R. M. Newman.

Lamiroy, B. (1987) "The Complementation of Aspectual Verbs in French," in Language Volume 63, Number 2, Pages 278-298.

Lebeaux, D. (1984) "Anaphoric Binding and the Definition of PRO' NELS 13

Manzini, M. R. (1983) "On Control and Control Theory" LI 14 # 3

Newman, P. and R. G. Schuh (1974) "The Hausa Aspect System,". Afroasiatic Linguistics, 1.1

Newmayer, F. J. (1975) English Aspectual Verbs. The Hague: Mouton.

Parsons, F. W. (1981) Writings on Hausa Grammar: The Collected Papers of F. W. Parsons", edited by G. L. Furniss.

Perlmutter, D. (1970) "The two verbs *begin*". Readings in English Transformational Grammar, ed. by R. Jacobs and P. Rosenbaum.

Pilszczikowa, N. (1960) "Les verbes auxiliaires en Haoussa". Rocznic Orientalistyczny 23(2):101-18.

Postal, P. (1970) On Raising. Cambridge, MA: MIT Press.

Rizzi, L. (1986) "Null objects in Italian and the Theory of Pro LI 17.3: 501-557.

Schuh, R. G. (1985) " Tense, Aspect, and Mood (TAM) System . MS, UCLA.

Tuller, L. (1982) "Null Subjects and Objects in Hausa." MS. UCLA.

Tuller, L. (1986) Bijection Relations in UG and the Syntax of Hausa. Ph.D. dissertation, UCLA.

Tuller, L. (1987) "AGR-drop and VP Focus-Fronting in Hausa . UQAM.

Yalwa, L. D. (1992) "Distribution of Pronouns and Anaphors in Hausa" M. A. dissertation, UCLA.

Yalwa, L. D. (in preparation) "Issues in the Syntax of Hausa Complementation". (Ph.D. dissertation, UCLA).

# DERIVING THE DISTRIBUTION OF CONJUNCTIONS

Ed Zoerner
University of California, Irvine

Abstract: This paper proposes an explanation for the limited possibilities of realized conjunctions in multitermed coordinations. It claims that conjunction (&) heads a fully articulated phrase (&P), which can iterate &P-shells, similar to V in Larson's VP-shell hypothesis. This structure enables a single & to unify any number of conjuncts (at LF), and thereby accounts for cases in which only a single overt & appears. Instances in which multiple &s surface are best understood as a PF condition on realizing emphasis.

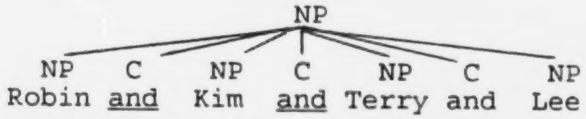
## The Puzzle

The distribution of overt conjunctions in multitermed coordinations is extremely limited. A conjunction must appear before (or cliticize to) the final conjunct, and may appear between each conjunct; no other options are possible:

1. a. Robin, Kim, Terry, and Lee
   b. Robin, and Kim, and Terry, and Lee
   c. *Robin, and Kim, Terry, Lee
   d. *Robin, Kim, and Terry, Lee
   e. *Robin, and Kim, and Terry, Lee

The form in (1a) presents the standard reading; (1b) places an emphasis on each conjunct. Most syntactic literature assumes that coordination has a flat structure, with a base-generated conjunction between each term, as in:

2.



In the standard view, the underlined forms represent those which can undergo Conjunction Reduction (CR) to yield the sequence of (1a). Should CR not apply, the emphatic reading of (1b) results. However, this account suffers from several problems. First, the structure in (2) clearly violates X-bar theoretical requirements, and does not conform to the general restriction on binary branching. Furthermore, the

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 217-233

process of CR is purely stipulative. Note that Japanese, a head-final language, also requires that a conjunction appear before the final conjunct (*to* 'and' coordinates NPs):

3. a. Robin, Kim, Terry to, Lee
   b. Robin to, Kim to, Terry to, Lee
   c. *Robin to, Kim, Terry, Lee
   d. *Robin, Kim to, Terry, Lee
   e. *Robin to, Kim to, Terry, Lee

No theoretical principles dictate that only the final conjunction should resist CR regardless of the head-parameter of the language.

The flat structure of coordination therefore fails. Any competing representation, however, must account for the following at first puzzling facts:

- A single lexically realized conjunction can unite more than two conjuncts.
- The occurrence of an overt conjunction between each conjunct provides one means of attaining emphasis.
- If there is a single conjunction, it must precede the rightmost conjunct, regardless of language.

This paper contains an analysis which derives the above facts through principled means, while remaining within the so-called Government-Binding tradition. Importantly, this work will adopt the Minimalist approach of Chomsky (1992), which recognizes only the representational levels of Phonetic Form (PF) and Logical Form (LF).

## Conjunction as a Syntactic Head

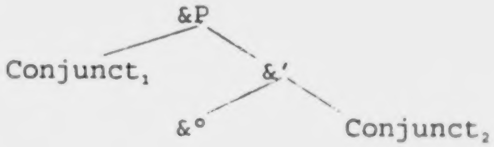
The analysis begins with the following claim: conjunction is a functional element (call it &) that heads its own fully articulated category (&P). This position is not entirely new; among others, Abney (1987), Collins (1988) and Munn (1992) have suggested the functional nature of &. Note that & possesses the following characteristics that Abney (1987: 64-5) holds indicative of functional elements:

- Functional elements constitute closed lexical classes.
- Functional elements are generally phonologically and morphologically dependent. They are generally stressless, often clitics or affixes, and sometimes even phonologically null.

- Functional elements lack..."descriptive content..." They mark grammatical or relational features, rather than picking out a class of objects.

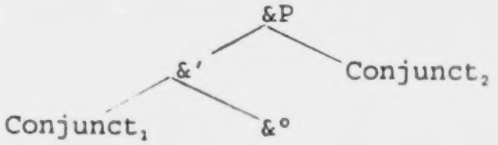
Given that & satisfies the above traits as do other functional elements such as Det, C, and Agr, it follows that & too should head a complete syntactic phrase, which projects its own specifier and complement positions:

4.



The above depicts a two-termed coordination in English; a head-final language such as Japanese presents a mirror-image:

5.



Assume that a conjunct enters a coordination by virtue of a structural relationship with an &° at LF; either a head-complement or specifier-head relation suffices to bring a conjunct within the coordination. Should a conjunct not stand in an appropriate relation to an &° at LF, the coordination is illformed.

Prosodic facts support the &P-structures represented here. As is well-known, in English the conjunction forms a prosodic unit with the second conjunct, while in Japanese the conjunction joins with the first conjunct (commas represent phrasal breaks):

6. a. Robin, and Kim
   b. *Robin and, Kim
   c. Robin to, Kim
   d. *Robin, to Kim

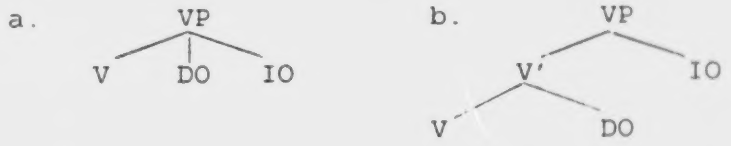
These facts fall out under the current proposal; in both languages the constituent of &' proves the relevant one to prosody. The flat structure of coordination cannot arrive at the facts as directly.

The &P-structure shown thus far provides a means of coordinating two conjuncts, but cannot yet represent multi-termed coordinations while still conforming to X-bar theoretical principles. An extension of Larson's (1988, 1990) VP-shell hypothesis applied to conjunction, however, enables this. A brief review of Larson's idea ensues, followed by its extension to the &P-structure.

The VP-Shell

Larson (1988) notes the inadequacy of previous representations of double-object constructions as in 'Robin sent a letter to Kim':

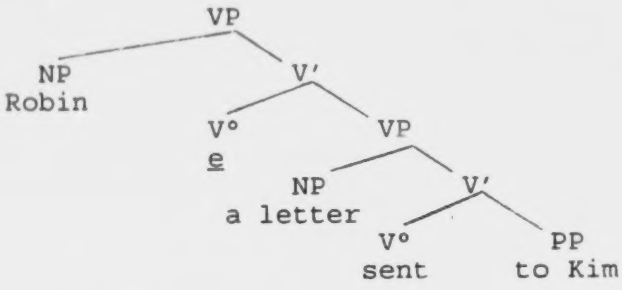
7.



The representation of (7a) contravenes the restriction to binary branching. In (7b), the indirect object stands apart from the verb and direct object; Larson (1988: 336-41) shows, though, that at some level the verb and indirect object should form a semantic/syntactic unit. As but one evidence, consider idiom chunks; in 'Robin threw Kim to the wolves', undeniably 'throw to the wolves' constitutes a single unit of meaning.

Larson therefore suggests a structure in which a V° position takes a VP as its complement, resulting in VP-shells. The underlying structure of 'Robin sent a letter to Kim' under this idea becomes:

8.

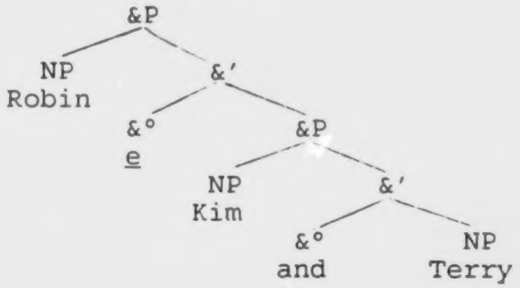


The verb raises to fill the underlyingly empty V° position for Case and agreement purposes, which yields the proper surface order. The verb together with the indirect object forms a constituent, that of V'; all branching is binary.

## Application to &P

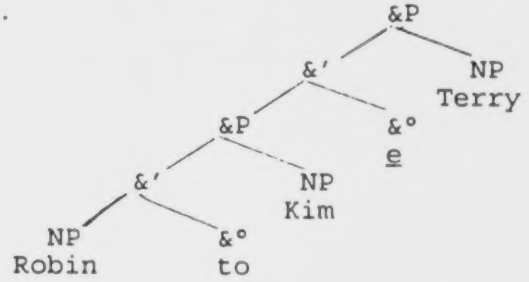
Since the VP-shell hypothesis does well in representing superficially ternary structures, it follows to apply a similar idea to coordination as well; differing only in not assuming an upper bound on the number of possible &P-shells. Doing so yields the following representation of a three-termed English coordination:

9.



Any coordination with *n* terms will have *n-1* &P nodes. The single lexical conjunction always occupies the lowest &° position. This holds in head-final languages as well:

10.



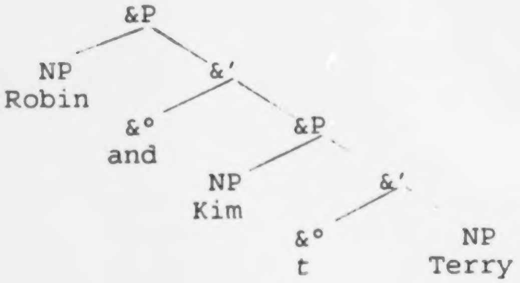
Adoption of the &P-shell hypothesis has two important consequences: It keeps coordination within binary branching restrictions, and depicts a coordination as having a single underlying conjunction. The desirability of the latter will become apparent in subsequent sections.

## &°-Movement

Recall the earlier assumption that conjuncts enter a coordination by virtue of their structural relationship with an &° term. An underlyingly empty &° position cannot conjoin a conjunct, just as in the VP-shell hypothesis an empty V° position cannot satisfy Case and agreement requirements. Coordinations with more than two terms, then, dictate movement of the &°$_i$

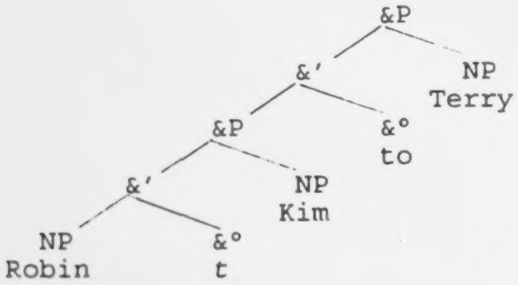
a Form-Chain operation (see Chomsky (1992: 21)) conjoins all terms appropriately. In English, this movement generally occurs at LF. After realization of the PF structure in (9), for example, the &° raises at LF so that it stands in an appropriate position to conjoin the highest conjunct through a specifier-head relation:

11.



In Japanese, an &° also undergoes Form-Chain to coordinate all conjuncts. Japanese differs from English, however, in that it requires this movement at PF instead of LF, for reasons to be explained shortly. The representation becomes:

12.



The movement shown above renders the correct surface order for phonetic interpretation; each conjunct stands in an appropriate structural relation with an &° for subsequent checking at LF. Both head-initial and head-final languages manage to bring all conjuncts into the coordination through structural relations with an &°, and differ only in at which level they do so. The first part of the puzzle now has an answer: A single conjunction can coordinate any number of terms by virtue of its chain-formation in &P-shells.

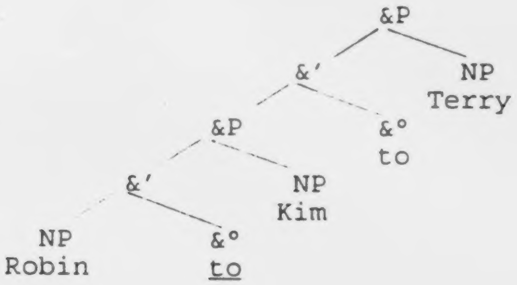
Realizing Emphasis

Clearly, phonological stress provides one means of creating an emphatic or focused reading; 'ROBIN went to the store' carries more emphasis than does 'Robin went to the store'. Assume, then, that emphatic readings result from PF phenomena.[1] Given this, the present analysis yields a straightforward account of how the grammar realizes emphatic coordinations such as 'Robin and Kim and Terry'. Specifically:

13. An emphatic reading of a coordination may result from the phonetic realization of the traces of the &°-chain at PF.

Consider first the Japanese case. The &° has left a trace in the lowest &° position at PF; when this trace copies the phonetic interpretation of its antecedent, the correct emphatic results (the underlined form shows the 'trace come to life'):

14.

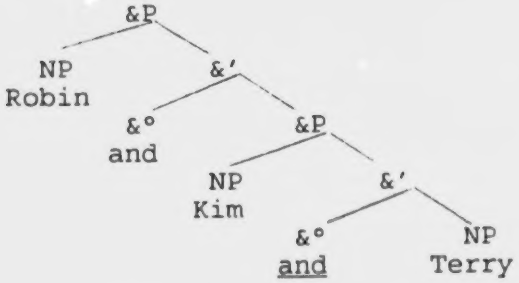


Because all traces in &°-positions result from a single Form-Chain operation, it follows that to attain the emphatic reading they will all assume phonetic content, regardless of the number of conjuncts. The analysis therefore correctly predicts (15a) as good, and the other forms of (15) as ungrammatical:

15. a. Robin to Kim to Terry to Lee
b. *Robin t Kim to Terry to Lee
c. *Robin to Kim t Terry to Lee

Now turn to English. As noted, in English, &°-raising generally occurs at LF. However, since emphasis is a PF phenomenon, reaching the emphatic reading of a coordination in English requires an earlier movement, at PF, instead. The following shows this result:

16.



The realized trace above occupies the position where the conjunction would stand in an non-emphasized coordination; the crucial difference between (9) and (16) is the presence of a lexical &° between the first two conjuncts in the latter, which contributes to the desired emphasis. Again, because &°-chain formation is a single operation, all traces assume content at once. The second aforementioned puzzling fact now has an explanation; the analysis correctly accounts for the fact that such emphatic readings require an overt conjunction between each conjunct through trace realization.

The analysis also correctly predicts that all such conjunctions will be phonetically identical, because the realized traces are in fact manifestations of the same single base-generated &°. For instance, from the movement as shown in (17a), (17b) but not (17c) can result to form an emphatic:

17. a. Robin and Terry t Kim t Lee
b. Robin and Terry and Kim and Lee
c. *Robin and Terry or Kim and Lee

Japanese provides an even stronger example. In Japanese both *to* and *ya* can conjoin NPs and translate as 'and'.[2] Note that they may not 'mix and match' to form an emphatic reading:

18. a. Robin to Kim to Terry to Lee
b. Robin ya Kim ya Terry ya Lee
c. *Robin ya Kim to Terry to Lee
d. *Robin to Kim ya Terry ya Lee

Theories of coordination which take each conjunction to be individually base-generated cannot arrive at the above facts without stipulation. This criticism applies to other efforts to bring conjunction within X-bar theory (e.g., Collins (1988)), and to any theoretical framework which assumes a phrase structure rule of the type: X → X Conj X.[3] Only the present analysis, which shows a single base-generated &° term per coordination, arrives at the facts straightforwardly.

### Support for the Analysis

The success of the above account hinges on the claim that a trace may assume phonetic content. Although this claim may lack precedent in the theory, it in fact has other useful applications and is not an ad hoc stipulation for coordination alone. Data from topicalization and passive and raising verb constructions demonstrate this.

Consider first topicalization. Assume that topicalization consists of movement of a phrase which adjoins to CP, as in:

19. $[_{TOP}$ Robin $[_{CP}$ Terry likes *t*]
'Robin, Terry likes'

Note now that should the trace assume the phonetic form of its antecedent, the following grammatical form results:

20. Robin, Terry likes Robin

This sentence indeed provides a natural way to emphasize that Robin in fact receives Terry's affection; perhaps the orthography 'Robin--Terry likes Robin!' indicates this more clearly.[4] Further similar examples include:

21. a. Peanuts--Robin asked for peanuts?!
b. Linguists--Kim can't stand linguists!
c. Lee--Terry thinks I like Lee?!

Notice that the underlined form must assume the same form as the topicalized element; generally, even a synonym fails:

22. ?*Peanuts--Robin asked for goobers?!

This suggests that the sentence-final element indeed is a copy of the moved topic rather than a random interjection.

Passive and raising verb data also support the idea of phonetically realized traces. These constructions involve movement of an NP from within the VP to the [SPEC, IP] position that leaves a trace:

23. a. $[_{IP}$ This story$_i$ $[_{I'}$ is $[_{VP}$ believed $t_i$ by all]]]
    b. $[_{IP}$ Robin$_i$ $[_{VP}$ seems $t_i$ to know the answer]]

Again, when the trace takes the phonetic form of its antecedent, a natural emphatic reading results:

24. a. This story is believed--this story!--by all
    b. Robin seems--Robin!--to know the answer

Notice that such an emphatic repetition of the subject may only occur in the location of the trace; other attempts crash:

25. a. *This story is--this story!--believed by all
    b. *This story is believed by--this story! --all
    c. *Robin seems to--Robin!--know the answer
    d. *Robin seems to know--Robin!--the answer

In sum, it appears that traces of argument-chains can indeed assume the phonetic interpretation of their antecedents, expressly for the purpose of attaining emphasis. The analysis of realizing emphatic coordinations therefore comes at no extra theoretical cost.

### &° as a Case Assigner

Recall the earlier claim that for nonemphatic readings, &°-movement in English takes place at LF, while in Japanese it occurs at PF. This distinction has the correct consequence of showing an overt &° as preceding the rightmost conjunct in all languages, but needs explanation; otherwise the present analysis will remain as stipulative as earlier CR accounts. In fact, given the &P-structure proposed here, the difference results from independent principles.

Consider: since &P is a maximal projection of a functional element, it blocks Case assignment from the verb to an NP in constructions such as the following:

26. $[_{\&P}$ Robin, Kim and Terry] left

The grammaticality of (26), however, indicates satisfaction of the Case Filter. The previous assumption seems to be that the verb assigns Case to the entire coordination:

27. $[_{\&P}$ Robin, Terry and Kim$]_{NOM}$ left

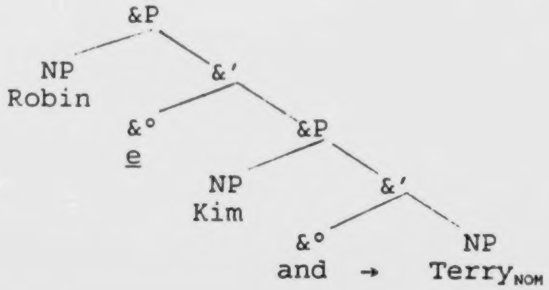
However, standard theory does not normally take functional categories as receiving Case. Furthermore, strictly speaking the Case Filter is not satisfied, inasmuch as an NP does not receive Case. An assumption more in keeping with the spirit of the Case Filter is that the final conjunct bears Case, as in:

28. $[_{\&P}$ Robin, Terry and Kim$_{NOM}]$ left

A single assignment of Case to an NP in a given position (here, [SPEC, IP]) manages to satisfy the Case Filter.[5] Since the verb cannot assign NOM through the functional &P nodes, the lexical &° must be assigning Case in (28). Along with verbs, then, conjunctions can assign Case, and like verbs, they may do so either through a head-complement or specifier-head relationship.

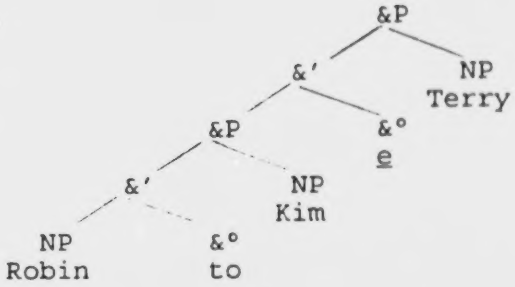
Assume that such &°-Case assignment in some sense incurs 'cost' on the grammar. If so, under Economy considerations it will only occur as a Last Resort, to salvage an otherwise Caseless construction. This explains why an overt conjunction does not raise in English until LF:

29.



Because the lexical &° can assign Case to the last (in linear sequence) NP from its base-generated position, the Case Filter is satisfied at PF, and the raising of the &° to coordinate all terms will not occur until LF in nonemphatic readings (such movement also being subject to Economy considerations, such as Procrastinate). Now turn to the base-generated Japanese construction:

30.



Here, Case assignment of the &° *to* to the NP 'Robin' through a head-complement relation, or to the NP 'Kim' through a specifier-head relation, would violate Economy. The rightmost NP 'Terry' here represents the Last Resort. Therefore, the &° must raise at PF to assign Case to 'Terry'. The Case Filter provides ample motive to override the Procrastinate nature of such &°-movement. All languages therefore realize an overt conjunction before the final NP in a multi-termed coordination for Case purposes, and the analysis therefore needs no stipulative devices such as the CR account does.

Although the notion of &° as a Case assigner may appear odd at first, data from English give it credence. Consider the following:

31. a. I left
b. *Me left
c. [Terry and I] left
d. [Terry and me] left
e. Robin gave him a nickel
f. *Robin gave he a nickel
g. Robin gave [Terry and him] a nickel
h. Robin gave [Terry and he] a nickel

No speaker ever utters (31b) instead of (31a), nor (31f) for (31e). Yet many speakers freely produce and accept both (31d) and (31h) instead of their prescriptively correct counterparts. This possibility follows from &°'s status as a Case assigner. The Case Filter merely requires that particular positions exhibit Case. Unlike a V°, an &° will not assign a particular Case by virtue of a particular structural relation. It can assign NOM, as it does to 'I' in (31c), for instance, or to 'he' in (31h). An &° can also assign Accusative Case, as it does to 'me' in (31d) and to 'him' in(31g). Although there may be a pervading prescriptive convention for the Case an &° assigns to match that which one normally expects of that position (e.g. NOM in a subject position), nothing in principle rules out a deviation from this; the &° assigns whichever case it desires, so to speak. The relatively acceptable digressions from the norm in (31) only result when one takes &° as a possible assigner of Case.[6]

Notice too that some languages, such as Chinese, require a lexical &° to coordinate NPs, but not to coordinate other phrases. That is:

32. a. *[Robin, Kim] mai-le yi-ben shu
buy-ASP one-CL book
'Robin, (and) Kim, bought a book'
b. [Robin gan Kim] mai-le yi-ben shu
'Robin and Kim bought a book'
c. Robin [chang qu, tiao wu]
sing song dance dance
'Robin sings a song, (and) dances a dance'

The analysis explains the ungrammaticality of (32a); it is illformed because no lexical &° assigns Case. Note that inserting the conjunction *gan* rescues the sentence. In (32c), a phonetically null &° may coordinate VPs because they do not need Case.

In coordinations in head-final languages of phrases other than NP, an overt conjunction still must precede the ultimate conjunct. For example, note the facts of Japanese coordinated sentences (Japanese *shi* conjoins CPs):

33. a. Robin-ga yomu, Kim-ga nomu shi,
SUB read SUB drink and
Terry-ga neru
SUB sleep
'Robin reads, Kim drinks, and Terry sleeps'
b. *Robin-ga yomu shi, Kim-ga nomu,
Terry-ga neru

(33b) crashes because no coordinator precedes the final clause. This fact naturally does not follow from Case considerations; presumably, though, the grammar prefers a consistent pattern of conjunction, and non-NP coordination must parallel NP-coordination. Though the intuition is clear enough, further investigation is needed to give it substance.

Summary

Recall that coordination phenomena presented three puzzling facts:

- A single lexically realized conjunction can unite more than two conjuncts.
- The occurrence of an overt conjunction between each conjunct provides one means of attaining emphasis.
- If there is a single conjunction, it must precede the rightmost conjunct, regardless of language type.

The analysis presented here arrives at the above facts through principled means. A single conjunction coordinates multiple conjuncts by virtue of its Form-Chain operation targeting all &° positions with its &P-shells. The emphatic reading results from the phonetic realization of all traces of this operation. The necessity of an overt conjunction preceding the final conjunct falls out directly from Economy principles once &° is taken to assign Case.

In short, the proposal of &° as a Case-assigning functional head of iterating maximal projections enables the correct prediction of the distribution of lexical conjunctions without resort to theoretically unmotivated stipulations; all the mechanisms used in the foregoing analysis are independently motivated apart from coordination facts. This success alone goes a long way in justifying the analysis; further research will no doubt uncover further theoretical and empirical benefits it confers.

## NOTES

Jim Huang, Naoki Fukui and especially Robert May deserve thanks for their helpful discussion with this paper. An anonymous reviewer contributed several useful suggestions of presentation. Hiromu Sakai kindly helped with the Japanese data, and Teresa Griffith graciously proofread an earlier version of the work. The author naturally assumes responsibility for remaining errors.

[1] This is admittedly a sketchy discussion of emphasis and focus; see Rochemont and Culicover (1990), for example, for further detail.

[2] Japanese *ya* and *to* differ slightly; roughly speaking, a speaker uses *ya* to suggest that the coordination may not be a closed set; the effect is similar to English 'among others, Robin and Kim...' The two coordinators are logically and truth-conditionally equivalent, however.

[3] For instance, Generalized Phrase Structure Grammar (GPSG) as in Sag et al (1985) and Categorial Grammar (CG) as in Steedman (1985) seem particularly susceptible to such a criticism. GPSG is required to posit two separate coordination schemata to account for what we have called the normal and emphatic readings; CG seems entirely unable to guarantee that all conjunctions in a multitermed coordination will be identical.

[4] An anonymous reviewer points out that standard left-dislocation constructions may give a more natural way to express emphasis:

i) Robin$_i$, Terry likes her$_i$

Nonetheless, the topicalized examples under discussion, though perhaps awkward for some speakers, provide a grammatical way to express emphasis or incredulity.

[5] This paper necessarily adopts an (admittedly tentative) interpretation of the Case Filter that differs somewhat from standard theory. It proceeds on the idea that the Case Filter applies to specific positions rather than NPs per se. For example, the [SPEC, IP] position must bear Case somewhere, as must a [COMP, V°] position, and so on. Under this idea, not every NP in (28) must have Case; one instance of Case-assignment satisfies the Case Filter requirements of the [SPEC, IP] position.

[6] Note that under the interpretation of the Case Filter in note (5), only one NP in a coordination needs to receive Case assignment. All other NPs, then, are in a sense free to select how they will surface. For example:

i) [Me and Robin$_{NOM}$] left

Under the present theory, 'and' assigns Case to 'Robin' and leaves the higher NP 'me' alone. Since this higher NP is not subject to any particular Case requirements, it may indeed surface with Accusative Case without contravening any principles, and the sentence as a result remains grammatical. An interesting question, though, is why the following should be so much worse than (i) above:

ii) *[I and Robin] left

Here, the higher NP 'I' actually bears NOM, which one expects to find under [SPEC, IP]; yet the form is bad. This fact poses a mystery for all theories of coordination, and needs further investigation.

The author welcomes readers to send comments or questions via e-mail to ezoerner@orion.oac.uci.edu.

## REFERENCES

Abney, Scott. 1987. The English noun phrase in its sentential aspect. MIT dissertation.

Chomsky, Noam. 1992. A minimalist program for linguistic theory. MIT occasional papers in linguistics, number 1. Cambridge MA: MIT Working Papers in Linguistics.

-----. 1991. Some notes on economy of derivation and representation. in Principles and parameters in comparative grammar, ed. by R. Freidin, 417-54. Cambridge MA: MIT Press.

Collins, Chris. 1988. Part I: conjunction adverbs. unpublished ms., MIT.

Gleitman, Lila. 1965. Coordinating conjunctions in English. Lg. 41.260-93.

Kayne, Richard. 1984. Connectedness and binary branching. Dordrecht: Foris.

Larson, Richard. 1990. Double objects revisited: reply to Jackendoff. LI 21.589-632.

-----. 1988. On the double object construction. LI 19.335-91.

Munn, Alan. 1992. A null operator analysis of ATB gaps. Ling. Rev. 9.1-26.

van Oirsouw, Robert. 1987. The syntax of coordination. New York: Croom Helm.

Rochemont, Michael and Peter Culicover. 1990. English focus constructions and the theory of grammar. Cambridge: Cambridge University Press.

Sag, Ivan, Gerald Gazdar, Thomas Wasow and Steven Weisler. 1985. Coordination and how to distinguish categories. NLLT 3.117-71.

Steedmanm Mark. 1985. Dependency and coordination in the grammar of English and Dutch. Lg. 61.523-68.

# L2 ACQUISITION OF ENGLISH REFLEXIVES BY NATIVE SPEAKERS OF KOREAN

Hye-Ryun Kim
University of Florida

Abstract: This study investigates how Korean adult learners of English at various levels of proficiency interpret English reflexives. The results of the experiment showed that Korean learners allowed the nonlocal antecedent about 35% of the time, which suggests that the Subset Principle does not operate in second language acquisition. However, the most advanced learners showed no performance difference from native speakers of English. This finding suggests that parameter resetting to the L2 value is possible.

## I. Introduction

One of the central issues in second language acquisition research concerns the question of whether UG is still accessible to the adult L2 learner. While researchers such as White (1988) and Flynn (1987) maintain that principles and parameters of UG are available to adult L2 learners, others (Bley-Vroman, 1989; Schachter, 1988; Clahsen and Muysken, 1986) argue that this is not the case.

In order to determine which position is indeed correct, we need to investigate whether or not L2 learners have linguistic knowledge for which there is no evidence in the input data. If L2 learners attain any type of knowledge which is attributable to UG, then we have good reason to believe that UG operates in adult L2 acquisition. However, if both the L1 and the L2 show the operation of a particular principle of UG, or if the L1 and the L2 have the same value for a particular parameter, although L2 learners demonstrate knowledge of the relevant properties of the L2, there is no way of knowing whether this knowledge is attributed to the availability of UG or to transfer of L1 knowledge. Therefore, to provide strong evidence for the operation of UG in adult L2 acquisition, it is necessary to eliminate effects of the L1. White (1990) suggest the following two situations where the effects of the L1 can be eliminated (p.128):

(1) a. some principle operates in the L2 but not the L1,

*Kansas Working Papers in Linguistics*, Volume 19, Number 1, 1994, pp. 235-261

and
b. the input underdetermines the L2 grammar.
(2) a. the L1 and the L2 have different values for some parameter, and
b. the input underdetermines the L2 grammar.

(1) is concerned with the operation of principles of UG. If L2 learners demonstrate knowledge of a principle of UG which is not instantiated in their L1, this will provide support for the claim that UG is still accessible to adult L2 learners. (2) is relevant to the parameters of UG. If L2 learners, under the situation (2), acquire the proper L2 value of a parameter, this will also be evidence for the operation of UG in adult L2 acquisition.

Second language acquisition of English reflexives by native speakers of Korean meets the conditions in (2). English and Korean differ in the domain in which reflexives may be bound and it is unlikely that the binding domain of English reflexives is explicitly taught in the classroom. In addition, the fact that English is a subset of Korean with respect to the reflexive binding provides an interesting research question about the Subset Principle. We may ask if the Subset Principle argued to operate in L1 acquisition also operates in L2 acquisition.

This paper reports on an experimental study which investigates how native speakers of Korean at various levels of English proficiency interpret English reflexives. The Governing Category Parameter and the Proper Antecedent Parameter of Wexler and Manzini (1987) are studied in relation to the Subset Principle. The issue of whether UG is still available to adult second language learners is discussed on the basis of results of the experiment.

## II. Binding Theory and Language Acquisition

1. The Governing Category Parameter and the Proper Antecedent Parameter. In relation to the binding theory, Wexler and Manzini (1987) propose the Governing Category Parameter (GCP) and the Proper Antecedent Parameter (PAP). Wexler and Manzini argue that since languages vary as to what counts as a governing category, the notion of governing category be parameterized as in (3):

(3) $\gamma$ is the governing category for $\alpha$ iff $\gamma$ is the minimal category that contains $\alpha$ and a governor for $\alpha$ and has
a) a subject; or

b) an INFL; or
c) a tense; or
d) a "referential" tense (=indicative mood); or
e) a "root" tense

This parameter is concerned with how far away the antecedent can be from the reflexive. In languages like English, reflexives must be bound within the same clause. However, in languages like Japanese and Korean, reflexives may take any NP as an antecedent as far as it is within the main clause. Languages like Russian treat reflexives differently depending on whether they occur in finite or nonfinite clauses. Consider the following sentence as an illustration:

(4) John thinks that [Bill wants [Tom to love himself]].

In (4), English reflexives allow only Tom to be an antecedent, Russian reflexives allow both Bill and Tom but not John as a potential antecedent, and Korean reflexives allow all three of them to be an antecedent.

Another parameter proposed by Wexler and Manzini (1987) is the Proper Antecedent Parameter. It is defined as in (5):

(5) A proper antecedent for $\alpha$ is
a. a subject $\beta$; or
b. any element $\beta$ whatsoever.

This parameter is concerned with what types of NPs can serve as antecedents for reflexives. While languages like English allow subjects and nonsubjects to be the antecedents of reflexives, languages like Korean and Japanese allow only subjects as the antecedents of reflexives. Consider the following sentence:

(6) John showed Bill a picture of himself.

In (6) in the case of English, both John and Bill can serve as the antecedent of the reflexive. On the other hand, in Korean or Japanese, only John can be the antecedent of the reflexive.

2. The Subset Principle and the Parameters of Binding Theory. The Governing Category Parameter and the Proper Antecedent Parameter are parameters of UG which meet the Subset Condition. That is, they yield languages which fall into a subset relation. This is illustrated as follows:

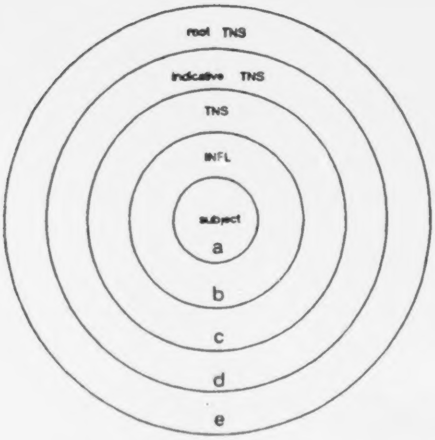


Figure 1. The governing category parameter

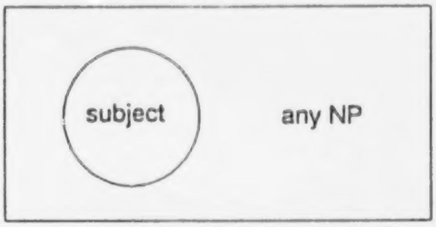


Figure 2. The proper antecedent parameter

As figure 1 indicates, the values of the GCP form a hierarchy. Languages like English which assume value (a) of the GCP are a subset of other languages which assume value (b), (c), (d) or (e) with reference to reflexive binding. Languages like Korean which assume value (e) of the GCP are a superset of other languages which assume value (a), (b), (c) or (d). Languages like Russian which assume value (c) are a subset of languages assuming value (d) or (e) but a superset of languages assuming value (a) or (b). Thus, as far as the GCP is concerned, English is one of the most restrictive languages and Korean is one of the least restrictive languages. With regard to the PAP, languages like Korean are a subset of languages like English since the latter allow any NPs as the antecedents of reflexives, whereas the former only allow subjects.

The Subset Principle states that given two languages, one of which is a subset of the other, if both are compatible with the input data, the learning function must pick the smaller one. Considering this principle from the developmental perspective, it is predicted that in actual language acquisition process, the child would start with the most restrictive values of the GCP and the PAP, and then switch to the appropriate L1 values, if there is evidence to the contrary. In terms of Korean learners' L2 acquisition of English reflexives, assuming the operation of the Subset Principle, it is expected that Korean learners would correctly bind English reflexives within the nearest clause from the beginning. With respect to the PAP, if the Subset Principle works in L2 acquisition, Korean learners of English would only allow subjects as the antecedents of English reflexives.

## III. Previous Studies

1. Finer and Broselow (1986). Finer and Broselow reported on a pilot study on the second language acquisition of English reflexives by six adult Koreans. Subjects were students in an intensive English program at a university in the USA. The experiment involved a picture identification task. Subjects were shown pairs of pictures and were given a sentence. They were then asked to indicate which of the two pictures was appropriate for the sentence they had heard or if both pictures represent the sentence. Examples of the types of the sentences used in the experiment are shown below:

(7) a. Mr. Fat thinks that Mr. Thin will paint himself.
    b. Mr. Thin asks Mr. Fat to paint himself.

Finer and Broselow's results are as in Table 1.

| | Local | Nonlocal | Either |
|---|---|---|---|
| Tensed Clause | 91.7 | 8.3 | 0 |
| Infinitive Clause | 58.3 | 37.5 | 4.2 |

Table 1. The Percentage of responses in tensed and infinitive clauses (Source: Finer and Broselow, 1986)

Table 1 shows that while subjects correctly chose the local antecedent in the tensed clauses (over 90% of the responses), they often failed to do so in the infinitive clauses. Finer and Broselow interpreted this result as indicating that subjects were assuming a value for the Governing Category Parameter which is somewhere between the (a) value (English) and the (e) value (Korean), possibly (c) (Russian). This is because the tensed/infinitival distinction is irrelevant to the distribution and interpretation of reflexive anaphors in both Korean and English. Finer and Broselow concluded that Korean learners of English had come up with a set of binding principles that is consistent with the parameters provided by UG, but inconsistent with either the first or the second language.[1]

Finer and Broselow provided another interpretation of the results in terms of the Subset Principle. The results from the tensed clauses support the operation of the Subset Principle. However, the results from the infinitive clauses are inconsistent with the operation of the Subset Principle. To account for the differences in responses between the tensed clauses and the infinitive clauses, Finer and Broselow suggested the possibility that Korean learners of English misanalyzed the subject of the infinitive as the direct object of the matrix verb. That is, Korean-speaking learners may consider the subject of the infinitive as the direct object of the matrix verb and avoid it as the antecedent because of the Proper Antecedent Parameter (the unmarked value of the Proper Antecedent Parameter is defined to be a subject). If this is indeed the case, it can be argued that the Subset Principle fully operates in second language acquisition: the learners, assuming the most unmarked value for the Governing Category Parameter and for the Proper Antecedent Parameter, preferred the local interpretation in the tensed clauses and the nonlocal interpretation in the infinitivals.

2. Thomas (1989). Thomas investigated whether second language learners know the following facts about the interpretation of English reflexives (p. 283):

(8) a. The Clause Mate Condition
In a finite clause, a reflexive pronoun must find an antecedent within the minimal dominating S-node.
b. The Subject Strategy
A reflexive is usually identified with a subject rather than a non-subject NP in an ambiguous context.

96 second language learners of English were presented with 30 sentences like those in (9) and were required to identify the antecedent of a reflexive pronoun by circling one of three multiple-choice answers. (For example, in case of sentence (9a), a subject may choose either (a) Bill, (b) David, or (c) Either Bill or David.)

(9) a. David could see that Bill was looking at himself in the mirror.
b. Mary angrily told me that Sue had spilled a lot of paint on herself.
c. Susan gave Mary three photographs of herself taken last summer.
d. After the medical test were completed, the doctor informed Bill about himself.

Sentences like (9a) and (9b) were constructed to investigate the subjects' knowledge of the clause mate condition in a coreference neutral context and a non-local antecedent pragmatically favored context, respectively. Sentences like (9c) and (9d) were presented to see if the subject strategy holds in a neutral context and a non-subject favored context, respectively.

The results of the experiment showed that, first, second language learners as a group do not fully obey the clause mate condition, regardless of the existence of pragmatic bias; second, second language learners and native speakers of English responded similarly to sentences like (9c) and (9d) by choosing the subject NP in a coreference neutral context and the non-subject NP in a non-subject NP favored context.

To see whether second language learners transfer the rules of L1 to L2, Thomas compared responses of native speakers of Spanish (29 subjects) with those of native speakers of Chinese (24 subjects) to the test sentences. Since Spanish allows only antecedents in the same clause, whereas Chinese allows non-local antecedents, it is expected that Chinese speakers would make more mistakes than Spanish speakers. However, no difference between Spanish and Chinese speakers' responses was found with respect to the clause mate condition. That is, both groups allowed long-distance binding of reflexives.. The following shows mean percentages of responses in bi-clausal finite sentences:

| | Spanish L1 (n=29) | Chinese L1 (n=24) |
|---|---|---|
| Non-local | 18.90 | 7.29 |
| Local | 59.48 | 69.04 |
| Either | 21.14 | 23.46 |

Table 2. Mean percentages of responses in bi-clausal finite sentences: coreference neutral (Source: Thomas, 1989)

Applying Wexler and Manzini's parameter-setting model of L1 acquisition to L2 acquisition, Thomas concluded that the Subset Principle does not operate in L2 acquisition, since the non-local responses are inconsistent with the subset value of the Governing Category Parameter. She also concluded that transfer of the grammar of L1 cannot account for the non-local responses of native speakers of Spanish, while this is possible for the Chinese. With respect to the Proper Antecedent Parameter, the fact that most of the subjects allowed both subject and non-subject antecedents in single clause sentences suggested that second language learners may successfully reset the Proper Antecedent Parameter to its marked value.

3. Hirakawa (1990). Hirakawa investigated how native speakers of Japanese learning English in Japan acquire syntactic properties of English reflexives and the effects of two parameters of UG, the Governing Category Parameter and the Proper Antecedent Parameter. With respect to the GCP, English represents the most unmarked value while Japanese represents the most marked one. Thus, if the Subset Principle operates in L2 acquisition, Japanese learners of English will correctly choose local antecedents in bi-clausal sentences by resetting the GCP back to its unmarked value. With reference to the PAP, if learners reset the parameter to its marked value, they will allow both subject and non-subject antecedents in single clausal sentences, since English represents the marked value in this case.

65 native speakers of Japanese in Grades 10,11,12, and 13 (ages 15 to 19 years) were tested with sentences like those in (10). Subjects were required to choose the antecedent of the reflexive in each sentence by circling one of a set of given choices.

(10) Type A: Two-clause sentence (finite)
John said that Bill hit himself.
Type B: Three-clause sentence (finite)
Mary remembers that June said that Alice blamed herself.
Type C: Two-clause sentence (infinite)
Mary asked Ann to introduce herself.
Type D: Three-clause sentence (infinite)
Ann knows that Mary told June not to hate herself.
Type E: One-clause sentence
Bob talked to Paul about himself.

Results showed that the Japanese learners of English could not set the value of the GCP correctly;

they set the widest value, allowing non-local antecedents for the reflexive even in tensed clauses. This could be explained in terms of transfer of the L1 value. Thus, the Subset Principle does not seem to operate in L2 acquisition. With respect to the difference between the finite and infinite clauses in responses, the finding is consistent with that of Finer and Broselow (1986). Reflexives in infinitial clauses received more non-local responses than reflexives in finite clauses, as seen in Table 3:

| | Finite (Type A) | Infinite (Type C) |
|---|---|---|
| Non-local | 17.13 | 36.45 |
| Local | 76.95 | 55.14 |
| Either | 5.92 | 7.79 |

Table 3. Mean percentages of responses in two types of clauses (Source: Hirakawa, 1990)

Finer and Broselow (1986) argue that L2 learners set the GCP to an intermediate value, allowing non-local antecedents in infinite clauses but disallowing them in finite clauses. Hirakawa's result is inconsistent with this, since her subjects made a large number of mistakes in Type A sentences.

With respect to the PAP, Hirakawa reported the similar results to those of previous studies. The Japanese learners of English correctly set the value of the PAP to the superset English value, allowing both subject and non-subject antecedents in one-clause sentences.

Although the majority of L2 learners failed to set the correct value of the GCP, some successful learners responded 100% correctly. On the basis of this, Hirakawa concluded that parameter resetting is possible, at least for some learners.

4. Summary. If the Subset Principle operates in L2 acquisition, second language learners of English should consistently choose the local antecedent of a reflexive. However, as we have seen, the results from Thomas (1989) and Hirakawa (1990) have not provided evidence for the operation of the Subset Principle in L2 acquisition.

The results of the studies also suggest that L2 learners' responses cannot be explained solely by L1 transfer. Spanish-speaking subjects in Thomas (1989) and

Korean-speaking subjects in Finer and Broselow (1986) did not treat English like Spanish and Korean, respectively.

As for the question of what value of the Governing Category Parameter L2 learners adopt, there is no agreement among studies. While Finer and Broselow's (1986) results suggest the intermediate value, the results of Thomas (1989) and Hirakawa (1990) suggest the widest value.

Another interesting question raised with reference to second language acquisition of English reflexives is whether UG is available to adult second language learners. The fact that the Subset Principle is not operative in L2 acquisition does not imply that UG is inaccessible to second language acquisition since the Subset Principle as a learning principle is independent of UG. As a matter of fact, the studies have shown that none of the subjects' responses was incompatible with a grammar of a natural language.[2]

As to whether parameter resetting is possible in second language acquisition, the studies do not provide a conclusive answer although Hirakawa (1990) suggested that "parameter resetting appears to be possible at least for some learners" (p. 81).

## IV. The Experimental Hypotheses

The research hypotheses can be stated as follows:

### The Governing Category Parameter

H1: Second language acquisition follows the course of first language acquisition. The Subset Principle operates in L2 acquisition, so Korean learners of English will adopt the subset L2 value immediately.

H2: The Subset Principle does not operate in L2 acquisition and the superset L1 value is transferred to the L2. Therefore, Korean learners of English will incorrectly allow long-distance antecedents.

H3: The Subset Principle does not operate and the L1 value is not transferred to the L2 either. Instead, an intermediate value is adopted as suggested by Finer and Broselow (1986). Then, the prediction is that learners will disallow the long-distance antecedent in tensed clauses but allow it in infinitival clauses.

H4: An unnatural possibility which violates UG is adopted. Learners may disallow the long-distance antecedent in infinitival clauses but allow it in tensed clauses.

The Proper Antecedent Parameter

H1: The Subset Principle operates or L1 transfer occurs,[3] so learners will start out with the unmarked value of the PAP, allowing only subjects as the antecedents of reflexives.

H2: Korean learners will immediately notice from the positive data from English that English has the marked value of the PAP. Therefore, learners will allow nonsubjects as well as subjects as the antecedents of reflexives.

H3: An unnatural grammar which violates UG will emerge. Learners may allow only nonsubjects as the antecedents of reflexives.

## V. The Experiment

1. Method. Subjects: The subjects tested in this experiment were 60 native speakers of Korean learning English, with 15 subjects in the each group. An English control group consisted of 15 native speakers of English attending University of Florida. Korean controls were 15 students who attended a college in Seoul, Korea. Background summary for each experimental group is as follows[4]:

Group 1 (n=15): this group comprised 9 female and 6 male grade 9 students aged 14 to 15. These subjects had studied English for 2 and half years at the time the test was administered. They received English lessons 4 times a week from nonnative instructors.

Group 2 (n=15): this group consisted of 7 female and 8 male grade 11 students aged from 16 to 17. These subjects had studied English for 4 and half years at the time of testing. They had English class 5 times a week from nonnative instructors. The Grammar-Translation Method was used for the purpose of helping students to prepare for the college entrance examination.

Group 3 (n=15): this group consisted of 8 female and 7 male first-year college students aged from 18 to 19. These subjects had studied English for 6 and half years at the time of testing. They had English class 5 times a week from nonnative instructors. 3 of 5 lessons were focused on reading comprehension and the other 2 lessons were focused on listening comprehension. 6 in this group of the subjects had experience with native instructors through private institutes.

Group 4 (n=15): this group consisted of 4 female and 11 male students who had studied in the USA for at least 3 years. All of them were graduate students at the University of Florida. Their ages ranged from 27 to 33. All these 15 had taken the TOEFL before first coming to the USA and scored above 550. Some of them had taken conversation courses at the University of Florida but none of them had been in the English Language Institute for intensive training.

Materials: Before the actual test, a pretest was given to students. The pretest was designed to test whether students knew the structures and vocabulary which would be used in the actual test. Students' knowledge of reflexive binding, i.e., the fact that reflexives cannot take extrasentential antecedents, was also tested. (The pretest is given in Appendix A.) Only the students who had correctly answered all the items of the pretest were included in the experimental groups. Of the high school students pretested, two third of the grade 9 students and a half of the grade 11 students failed. Of the college students pretested, only 3 failed, and none in group 4 failed in the pretest.

In the actual test, four different types of sentences were included, with 4 sentences of each type. Thus, a total 16 sentences were tested. The following presents the sentence types used in this study[5]:

(11) Type A: two-clause sentence containing a tensed embedded clause
John thinks that Tom likes himself.
Type B: two-clause sentence containing an infinitival embedded clause
John wants Tom to wash himself.
Type C: one-clause sentence
John told Tom about himself.
Type D: three-clause sentence containing an infinitival embedded clause
John thinks that Tom wants Bill to wash himself.

Types A and B sentences are relevant to the investigation of the GCP. Since the previous studies (Finer and Broselow, 1986; Hirakawa, 1990) found that reflexives in infinitival clauses had received more non-local responses than reflexives in tensed clauses, two types of sentence (A and B) were included for the purpose of comparison. Type C sentences were included to investigate the PAP. Type D sentences aimed to examine the GCP further. Finer and Broselow (1986) argued that their subjects were assuming an intermediate value of the GCP, distinguishing between tensed and infinitival clauses. If this was indeed the case, Korean learners should not choose the subjects of matrix clauses as the antecedents of reflexives in sentences like Type D, as suggested by Hirakawa (1990).

A Korean-speaking control group received a Korean version of the test. Each test sentence was translated into Korean as naturally as possible.[6]

Procedure: The test sentences were randomized and counterbalanced and were presented in the same order. The test was administered either individually or to a class. Before administering the test, it was determined that all the subjects knew that John, Tom and Bill were male names and Susan, Mary and Alice were female names. The subjects were informed that there was no time limit. However, they were instructed to answer each item based on how they felt about each sentence, without thinking too much. They were also instructed not to change their answers once they were done. Both written and oral instructions were given in Korean (For English controls, those instructions were given in English). The subjects were asked to indicate the antecedent of the reflexive by circling one of the given choices. Consider the following examples:

(12) John thinks that Tom likes himself.
    a. John
    b. Tom
    c. either John or Tom
    d. I don't know
(13) John thinks that Tom wants Bill to wash himself.
    a. John
    b. Tom
    c. Bill
    d. either John or Tom
    e. either Tom or Bill
    f. either John or Bill
    g. either John or Tom or Bill
    h. I don't know

In (12), the (a) and (b) answer choices refer to the potential antecedent appearing first and second in the sentence, that is, NP1 (John) and NP2 (Tom), respectively. The third answer, either NP1 and NP2, refers to the ambiguous interpretation of the sentence. For the case where the subject cannot decide the appropriate antecedent of the reflexive, "I don't know" answer was included as the fourth choice. In (13), in addition to the NP1 (John), NP2 (Tom) and NP3 (Bill), all the possible combinations of these three potential antecedents were given.

2. Results. Performance of control groups: Table 4 shows overall responses from the English control group and the Korean control group:

| | | English (n=15) | Korean (n=15) |
|---|---|---|---|
| Type A | NP1 | 0 (0.00%) | 29 (48.33%) |
| | NP2 | 60 (100.00%) | 19 (31.67%) |
| | NP1/2 | 0 (0.00%) | 12 (20.00%) |
| Type B | NP1 | 1 (1.67%) | 25 (41.67%) |
| | NP2 | 59 (98.33%) | 18 (30.00%) |
| | NP1/2 | 0 (0.00%) | 17 (28.33%) |
| Type C | NP1 | 41 (68.33%) | 55 (91.67%) |
| | NP2 | 13 (21.67%) | 0 (0.00%) |
| | NP1/2 | 6 (10.00%) | 5 (8.33%) |
| Type D | NP1 | 0 (0.00%) | 5 (8.33%) |
| | NP2 | 0 (0.00%) | 30 (50.00%) |
| | NP3 | 60 (100.00%) | 14 (23.33%) |
| | NP1/2 | 0 (0.00%) | 1 (1.67%) |
| | NP1/3 | 0 (0.00%) | 0 (0.00%) |
| | NP2/3 | 0 (0.00%) | 9 (15.00%) |
| | NP1/2/3 | 0 (0.00%) | 1 (1.67%) |

Table 4. Responses from the control groups

As expected, the English control group performed nearly perfectly with respect to the GCP. Regarding the PAP, native speakers of English allowed nonsubject as well as subject as antecedents of reflexives, with preference of subject to nonsubject antecedents, 68% to 22%. This is in contrast with the responses from the Korean control group. None of the Korean controls chose nonsubject antecedents for reflexives. With respect to the GCP, while native speakers of Korean showed all three types of responses (i.e., "NP1", "NP2", or "either NP1 or NP2") as expected, long-distance antecedents were most preferred.

Overall performance of experimental groups: Table 5 displays the number and the percentage of responses given by four experimental groups to four different types of sentences. Table 5 shows that all groups performed quite well across four types of sentences. With respect to Types A, B and D sentences, which are of relevance to the investigation of the GCP, the most frequent response

| | | Group 1 Number | ( n= 15 ) % | Group 2 Number | ( n= 15 ) % | Group 3 Number | ( n= 15 ) % | Group 4 Number | ( n= 15 ) % | Total Number | ( n= 60 ) % |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Type A | NP1 | 25 | 41.67 | 19 | 31.67 | 18 | 30.00 | 6 | 10.00 | 68 | 28.33 |
| | NP2 | 31 | 51.67 | 34 | 56.67 | 41 | 68.33 | 54 | 90.00 | 160 | 66.67 |
| | NP1/2 | 2 | 3.33 | 1 | 1.67 | 1 | 1.67 | 0 | 0.00 | 4 | 1.67 |
| | Dont' know | 2 | 3.33 | 6 | 10.00 | 0 | 0.00 | 0 | 0.00 | 8 | 3.33 |
| | Total | 60 | | 60 | | 60 | | 60 | | 240 | |
| Type B | NP1 | 18 | 30.00 | 18 | 30.00 | 12 | 20.00 | 5 | 8.33 | 53 | 22.08 |
| | NP2 | 39 | 65.00 | 34 | 56.67 | 45 | 75.00 | 55 | 91.67 | 173 | 72.08 |
| | NP1/2 | 2 | 3.33 | 0 | 0.00 | 2 | 3.33 | 0 | 0.00 | 4 | 1.67 |
| | Dont' know | 1 | 1.67 | 8 | 13.33 | 1 | 1.67 | 0 | 0.00 | 10 | 4.17 |
| | Total | 60 | | 60 | | 60 | | 60 | | 240 | |
| Type C | NP1 | 48 | 80.00 | 43 | 71.67 | 49 | 81.67 | 54 | 90.00 | 194 | 80.83 |
| | NP2 | 9 | 15.00 | 9 | 15.00 | 8 | 13.33 | 5 | 8.33 | 31 | 12.92 |
| | NP1/2 | 2 | 3.33 | 1 | 1.67 | 2 | 3.33 | 1 | 1.67 | 6 | 2.50 |
| | Dont' know | 1 | 1.67 | 7 | 11.67 | 1 | 1.67 | 0 | 0.00 | 9 | 3.75 |
| | Total | 60 | | 60 | | 60 | | 60 | | 240 | |
| Type D | NP1 | 4 | 6.67 | 1 | 1.67 | 4 | 6.67 | 0 | 0.00 | 9 | 3.75 |
| | NP2 | 12 | 20.00 | 19 | 31.67 | 10 | 16.67 | 5 | 8.33 | 46 | 19.17 |
| | NP3 | 34 | 56.67 | 35 | 58.33 | 38 | 63.33 | 55 | 91.67 | 162 | 67.50 |
| | NP1/2 | 3 | 5.00 | 2 | 3.33 | 0 | 0.00 | 0 | 0.00 | 5 | 2.08 |
| | NP1/3 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 |
| | NP2/3 | 5 | 8.33 | 0 | 0.00 | 5 | 8.33 | 0 | 0.00 | 10 | 4.17 |
| | NP1/2/3 | 1 | 1.67 | 2 | 3.33 | 3 | 5.00 | 0 | 0.00 | 6 | 2.50 |
| | Don't know | 1 | 1.67 | 1 | 1.67 | 0 | 0.00 | 0 | 0.00 | 2 | 0.83 |
| | Total | 60 | | 60 | | 60 | | 60 | | 240 | |

Table 5. The number and percentage of responses by experimental groups

was the local interpretation of the reflexives. Korean learners of English as a whole chose the local antecedent more than 65% of the time (66.67%, 72.08%, 67.50% for Types A, B and D, respectively). However, the long-distance antecedent was also chosen more than 20% of the time (28.33%, 22.08%, 22.92% for Types A, B and D, respectively), which means that the subjects, all together, had not yet acquired the behavior of English reflexives. (The analysis of individual subjects' responses revealed that this response pattern was consistent across subjects, with the exception of Group 4 subjects. While none of the subjects from Groups 1, 2 and 3 responded perfectly to types A, B and D sentences, 7 out of 15 subjects from Group 4 showed 100% correct responses to these sentences.) The "either NP1 or NP2" or "either NP1 or NP2 or NP3" responses were quite rare. The subjects tended not to judge the sentences to be ambiguous. The fact that the subjects showed nonlocal responses about 35% of the time suggests that the Subset Principle does not operate in L2 acquisition, since the Subset Principle predicts that subjects will only choose local antecedents for the reflexive. With respect to the PAP, the subjects correctly allowed nonsubjects as well as subjects as the antecedents of reflexives. This result indicates that the subjects had already reset the PAP to the marked value.

In order to determine whether there are significant differences in performance across groups and sentence types, two-way analysis of variance was done. Results of the ANOVA indicated a significant group effect (F(3,56)=14.71, p<.001) but no significant effect of sentence types (F(2,112)=1.29, p=.279) or interaction between groups and sentence types (F(6,112)=.60, p=.729). To refine the analysis, the data were further analyzed. Tukey's procedure revealed that Groups 1, 2 and 3 are significantly different from Group 4 (p<.05). Between Groups 1 and 2, 1 and 3, 2 and 3, no significant difference was found.

Performance of individual subjects: Thomas (1991) and Eckman (1993) argue that in order to determine whether L2 learners' grammars are constrained by UG, individual subjects' patterns of response should be investigated. Therefore, I present the number of subjects within each group who systematically bound reflexives to the indicated antecedent(s). In Table 6, 'systematicity' is defined as 3 or 4 identical responses to the 4 tokens of each sentence type.

Table 6 indicates that as subjects' proficiency

| | G1<br>(n=15) | G2<br>(n=15) | G3<br>(n=15) | G4<br>(n=15) |
|---|---|---|---|---|
| Type A | | | | |
| NP1 | 1 | 3 | 1 | 0 |
| NP2 | 4 | 7 | 11 | 14 |
| NP1/2 | 0 | 0 | 0 | 0 |
| Type B | | | | |
| NP1 | 1 | 1 | 1 | 0 |
| NP2 | 8 | 4 | 11 | 14 |
| NP1/2 | 0 | 0 | 0 | 0 |
| Type C | | | | |
| NP1 | 13 | 11 | 13 | 15 |
| NP2 | 0 | 0 | 0 | 0 |
| NP1/2 | 0 | 0 | 0 | 0 |
| Type D | | | | |
| NP1 | 0 | 0 | 1 | 0 |
| NP2 | 0 | 1 | 0 | 0 |
| NP3 | 7 | 6 | 9 | 14 |
| NP1/2 | 0 | 0 | 0 | 0 |
| NP1/3 | 0 | 0 | 0 | 0 |
| NP2/3 | 0 | 0 | 1 | 0 |
| NP1/2/3 | 0 | 0 | 0 | 0 |

Table 6. The number of subjects within each group who systematically bound reflexives to the indicated antecedent(s)

level increased, more subjects systematically bound reflexives only to local antecedents. In the case of Type C sentences, most subjects systematically bound reflexives only to subject NPs. The fact that only 1 out of 60 subjects systematically allowed ambiguous interpretations of reflexives (in Type D sentences) suggests that subjects habitually prefer one interpretation over the other(s).

Thomas (1991) and Eckman (1993) suggest that the response pattern which allows long-distance binding while disallowing local binding is not admitted by Wexler and Manzini's GCP. However, if we consider that native-speaker controls as well as L2 learners tend to avoid ambiguous interpretations, some subjects' responses allowing only long-distance binding can be explained as their preference for long-distance antecedents. Since long-distance antecedents are preferred to local antecedents in Korean, some subjects' systematic preference for long-distance binding may be due to L1 interference. The fact that 14 out of 15 subjects from group 4 systematically allowed only local binding, despite the preference for long-distance antecedents in Korean, suggests that these learners had already acquired the behavior of English reflexives.[7]

Performance on the governing category parameter: The results from the experimental groups showed that the subjects did not perform differently with respect to the Types A and B. This result was rather unexpected, since the previous studies found the difference in choice of local antecedents between sentences involving tensed embedded clauses and those involving infinitival embedded clauses. The results from Type D sentences confirmed no performance difference between tensed and infinitival clauses. The subjects correctly chose the local antecedent 67% and 68% of the time on Type A sentences and Type D sentences, respectively. Because no significant difference between three types of sentences was obtained, it was decided that the results from Types A, B and D were taken together. Table 7 displays the mean number and percentage of correct responses from each experimental group. Since total 12 sentences related to the GCP were included in the test, a perfect score is 12.

| | Group 1 | Group 2 | Group 3 | Group 4 |
|---|---|---|---|---|
| Score | 6.93<br>(57.75%) | 6.87<br>(57.25%) | 8.27<br>(68.92%) | 10.93<br>(91.08%) |

Table 7. Mean number and percentage of correct responses (by group)

Table 7 suggests that number of years of English study has an effect on the subjects' ability to correctly identify the antecedents of reflexives. Figure 3 displays the relationship between performance on the test and years of English study. As figure 3 indicates, there exists a positive relationship between years of English study and test score, with the exception of Group 2.

However, the fact that the Group 4 performed significantly better than the Groups 1, 2 and 3 suggests that the performance differences among four groups might be attributed to the amount of natural linguistic input each group had received. Group 4 had received lots of natural L2 input and had experience using English in communicative situations. In contrast, Groups 1 and 2 had

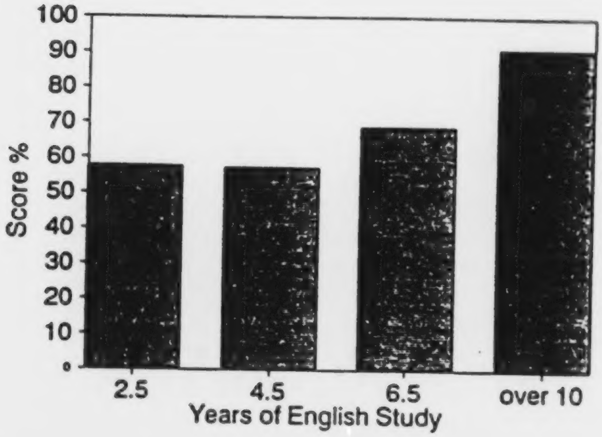


Figure 3. Relationship of score to number of years of English study

received no English training from native instructors and they had been exposed to English only in formal classroom settings. Most of the class time in these two groups were devoted to grammar instruction and translation. Group 3 also had had English training in formal classroom settings, but listening comprehension was an important part of English lessons and some of group 3 subjects had had experience with native instructors (by taking English conversation courses).

In order to determine whether there is a significant difference in performance between Group 4 and the English controp group, analysis of variance was done. The result showed no significant difference between two groups (F(1,28)=5.55, p<.05). Mean percentage of correct responses given by Group 4 and the English control group to Types A, B and D sentences is 91.08 and 99.44, respectively.

Performance on the proper antecedent parameter: The results from Type C sentences showed that Korean learners had already reset the PAP to its L2 value. Even the youngest group of the experimental groups correctly allowed nonsubjects as well as subjects as the antecedents of reflexives. Since only subjects are allowed to be the antecedents of reflexives in Korean, my subjects could not transfer the L1 value of the PAP to L2. This result raises questions about Finer and Broselow's (1986) interpretation that Korean L2 learners seemed to analyze the subject of the infinitive as the direct object of the matrix verb and to avoid it as the antecedent of the reflexive because of the PAP. According to Finer and Broselow, Korean L2 learners might assume the unmarked value of the PAP because of the operation of the Subset Principle and this might affect learners' performance on Type B sentences. However, the results from Type C sentences shows that Korean learners were not assuming the unmarked value of the PAP.

Although either subjects or nonsubjects are allowed as the antecedents of reflexives, Korean learners preferred subject to nonsubject antecedents (81% to 13%) as did English controls. However, L2 learners' response pattern did not replicate that of English controls. In fact, there was a significant difference between Korean learners and English controls ($\chi^2=9.916$, $p<.01$). The results for individual sentences revealed that Korean learners and English controls had not responded similarly to some sentences. For example, in the following sentence, where English controls chose the nonsubject antecedent almost exclusively, Korean learners most frequently chose the subject antecedent:

(14) Susan asked Mary about herself.
NP1 NP2

| | G1 (n=15) | G2 (n=15) | G3 (n=15) | G4 (n=15) | Control (n=15) |
|---|---|---|---|---|---|
| NP1 | 8 | 6 | 10 | 10 | 0 |
| NP2 | 6 | 6 | 4 | 4 | 13 |
| NP1/2 | 1 | 0 | 0 | 1 | 2 |
| Don't know | 0 | 3 | 1 | 0 | 0 |

Table 8. The number of responses to (14) by groups

Zobl (1989) argued that L2 learners show differential sensitivity to input properties: while L2 learners are sensitive to abstract properties of the L2 syntax, they are not very sensitive to discourse-pragmatic properties of the L2. The results from Type C sentences are consistent with this claim. While Korean learners reset the PAP to its L2 value, they do not seem to have semantic or pragmatic knowledge that native speakers of English have. For example, when Korean learners encounter sentences like (14), most of them, even the most advanced learners, will interpret the reflexive as referring to the subject, which is inconsistent with the interpretation that most native speakers of English assume.

## VI. Discussion

The results of this experiment suggest that the Subset Principle does not operate in L2 acquisition. If the Subset Principle were operating, Korean learners of English should consistently choose the local antecedents of reflexives. Since this proved not to be the case, we may conclude that the Subset Principle, a learning principle assumed to operate in L1 acquisition, does not continue to function in L2 acquisition.

The results from Types A and B sentences are inconsistent with those reported by Finer and Broselow (1986) and by Hirakawa (1990). Finer and Broselow found that there were more nonlocal responses for infinitival clauses than for tensed clauses. This finding was replicated by Hirakawa. In order to account for the differential performance on these two types of sentences, Finer and Broselow suggested that L2 learners might adopt an intermediate value of the GCP. Hirakawa also suggested that L2 learners might pass through the stage in which they assume an intermediate value of the GCP. This claim is disconfirmed in this experiment, since the subjects in the present experiment, at all levels of English proficiency, did not perform differently with respect to Types A and B sentences. To prove the hypothesis that learners move from the widest value to the narrower values as they become more proficient in English, further empirical investigation is required.

With respect to the issue of the availability of UG, the results of this experiment are consistent with White (1990)'s claim: although L1 transfer occurs initially, parameter resetting is possible, given appropriate input. The fact that the subjects from Groups 1, 2 and 3 allowed the long-distance antecedents of reflexives over 20% of the time suggests that they were transferring the superset L1 value of the GCP into their L2. However, the fact that the Group 4 subjects did not perform differently from the English controls suggest that if enough positive data from the L2 are given, it is possible to change from the superset L1 value to the subset L2 value. The ability to change from the L1 value of a parameter to the L2 value shows that UG must still be operating, since the parameter values are part of UG (White, 1989). Therefore, we may conclude that L2 acquisition as well as L1 acquisition is constrained by UG, at least in the domain tested.

## VII. Implications of the Results of the Experiment for L2 Pedagogy

The results of this experiment suggest that the natural linguistic input from the L2 may be critical to parameter resetting. The most advanced group of four experimental groups successfully reset the GCP to its L2 value: the learners performed in the same way as the native speakers of English. This group had been in an English-speaking environment for at least three years, using English in communicative situations. In contrast to this group, the other three groups had not received much natural input from native English speakers and had been exposed to English only in classroom settings. Based on these observations, we may argue that a large amount of positive data from the L2 is essential for parameter resetting in the case in which the L1 has the superset value and the L2 has the subset value of a parameter.

In the case of second language acquisition, as in the case of learning English in the USA, the positive data from the L2 will be sufficient for parameter resetting, at least in the case of the GCP. Hyams (1986) argued that during the developmental process of L1, the child's perception of the input data changes: there is some perceptual mechanism that filters out the relevant input data at the early stage, but brings them to the child's attention at some later stage. As in L1 acquisition, in L2 acquisition the learner's perception of the input data may also change. As the learner's level of proficiency in English goes up by getting more positive data from English, he or she may come to attend to more subtle aspects of English. The behavior of English reflexives is one such aspect. Thus, it may be that only advanced L2 learners are able to attend to and eventually acquire the properties of Engilsh reflexives, thereby resetting the GCP to its L2 value.

In the case of foreign language acquisition, as in the case of learning English in Korea, it is hard for learners to get enough positive input from English, since they are exposed to English only in formal classroom settings. In the situations like this, negative evidence[8] may be useful. In order to inform L2 learners that certain interpretations are impossible in English, several methods can be used. In addition to explicit grammar teaching, Rutherford and Sharwood Smith (1987)'s consciousness-raising (the deliberate attempt to draw the learner's attention specifically to the formal properties of the target language), or Tomasello and Herron (1990)'s garden-path method (whereby L2 learners are induced to make errors and then they are given corrections) may be useful. If L2 learners are not informed of the behavior of English reflexives and enough positive input is not given either, L2 learners seem to assume the superset value of the GCP and allow nonnativelike interpretations in sentences like Types A, B and D.

## NOTES

* An earlier version of this paper was presented at the conference on second language acquisition and pedagogy, October 1993, in Milwaukee, Wisconsin. I would like to express my appreciation to L. White, B. Schwartz, H. Zobl and G. Iverson for their helpful comments and suggestions.

[1] The interpretation similar to this was also presented in Finer (1989), and Broselow and Finer (1991): at a certain stage of acquisition, L2 learners seem to adopt a value of the GCP that is midway between the L1 and the L2 values.

[2] More recently, Thomas (1991) and Eckman (1993) reported some cases in which interlanguage grammars do not adhere to the constraints of UG.

[3] Since Korean represents the subset value of the PAP, L1 transfer results in the same effect as the operation of the Subset Principle. In cases like this, there is no way of knowing whether learners' behavior is attributable to the operation of the Subset Principle or to L1 transfer.

[4] Those who had been exposed to English before age 12, or those who had lived abroad before age 23 were not included in the experimental groups. Those who had had experience in linguistics were also excluded. Before administering the test, it was ascertained that none of the subjects had been explicitly taught the behavior of English reflexives.

[5] English reflexives are first taught in the classroom in grade 8. Since grade 9 students do not have a large vocabulary in English, the test sentences were

designed to include only the words which had been introduced in their English class. In the same vein, structures as simple as possible were used in the test. In order to avoid monotony, negative sentences as well as affirmative sentences were included, as in Hirakawa (1990).

[6] Fifteen Korean-speaking adults who achieved an advanced level of proficiency in English were asked to translate the test sentences into Korean. The most common interpretation for each sentence was used in the Korean controls.

[7] In order to determine whether L2 learners' responses represent their preference for one interpretation over the other(s) or their underlying grammars, a new methodology must be used. That is, we need to elicit L2 learners' judgment of different interpretations of the same sentence by presenting the learner with one context and one sentence at a time.

[8] Schwartz (1987), and Schwartz and Gubala-Ryzak (1992) argue that negative evidence cannot be used in L2 acquisition. However, according to Birdsong (1989), negative evidence is necessary and useful for the disconfirmation of certain hypothesis types in L2 acquisition.

## REFERENCES

Birdsong, D. 1989. Metalinguistic Performance and Interlinguistic Competence. New York: Springer.

Bley-Vroman, R. 1989. What is the Logical Problem of Foreign Language Learning? In S. Gass and J. Schachter (eds), Linguistic Perspectives on Second Language Acquisition. Cambridge: Cambridge University Press.

Broselow, E. and Finer D. 1991. Parameter Setting in Second Language Phonology and Syntax. Second Language Research.7.35-59.

Chomsky, N. 1981. Lectures on Government and Binding. Dordrecht: Foris.

Clahsen, H. and Muysken, P. 1986. The Availability of Universal Grammar to Adult and Child Learners: A Study of the Acquisition of German Word Order. Second Language Research.2.93-119.

Eckman, F. 1993. Local and Long-Distance Anaphora in Second Language Acquisition. Paper presented at the annual meeting of the American Association for Applied Linguistics, Atlanta, April 1993.

Finer, D. 1991. Binding Parameters in Second Language Acquisition. In L. Eubank (ed.), Point/Counterpoint: Universal Grammar in the Second Language. Amsterdam: John Benjamins.

Finer, D. and Broselow, E. 1986. Second Language Acquisition of Reflexive Binding. NELS.16.154-168.

Flynn, S. 1987. A Parameter-Setting Model of L2 Acquisition. Dordrecht: Reidel.

Hirakawa, M. 1990. A Study of the L2 Acquisition of English Reflexives. Second Language Research.6.60-85.

Rutherford, W. and Sharwood Smith, M. 1985. Consciousness-raising and Universal Grammar. Applied Linguistics.6.274-282.

Schachter, J. 1988. Second Language Acquisition and its Relationship to Universal Grammar. Applied Linguistics.9.219-235.

Schwartz, B. 1987. Ruling out Negative Evidence in Second Language Acquisition. ms. University of Southern California.

Schwartz, B. and Gubala-Ryzak, M. 1992. Learnability and Grammar Reorganization in L2A: Against Negative Evidence Causing the Unlearning of Verb Movement. Second Language Research.8.1-38.

Thomas, M. 1989. The Interpretation of English Reflexive Pronouns by Non-Native Speakers. Studies in Second Language Acquisition.11.281-301.

Thomas, M. 1991. Universal Grammar and the Interpretation of Reflexives in a Second Language. Language.67.211-239.

Tomasello, M. and Herron, C. 1988. Down the Garden Path: Inducing and Correcting Overgeneralization Errors in the Foreign Language Classroom. Applied Linguistics. 9.237-246.

Wexler, K. and Manzini, R. 1987. Parameters and Learnability in Binding Theory. In T. Roeper and E. Williams (eds), Parameter Setting. Dordrecht: Reidel.

White, L. 1988. Universal Grammar and Language Transfer. In J. Pankhurst, M. Sharwood Smith and P. Van Buren (eds), Learnability and Second Languages: A Book of Readings. Dordrecht: Foris.

White, L. 1990. Second Language Acquisition and Universal Grammar. Studies in Second Language Acquisition.12. 121-133.

Zobl, H. 1989. Modularity in Adult L2 Acquisition. Language Learning.39.29-79.

## APPENDICES

### Appendix A: Pretest

I. Vocabulary

| | | |
|---|---|---|
| hurt | respect | believe |
| picture | hate | wash |

II. Translation
1. John said that Tom played tennis.
2. John told Tom about the movie.
3. Susan told Mary to study English.
4. Susan wants Mary not to go swimming.
5. John thinks that Tom wants Bill to visit Mary.
6. Susan does not like Mary.

III. Reflexive-binding

In the following sentence who does himself refer to?

John and Tom were under the tree.
Tom was painting himself.
a. John B. Tom c. either John or Tom d. I don't know

### Appendix B: Test Sentences

Type A: two-clause sentences containing tensed embedded clauses

a. John thinks that Tom likes himself.

b. Susan does not believe that Mary hurt herself.
c. John said that Tom washed himself.
d. Susan knows that Mary does not respect herself.

Type B: two-clause sentences containing infinitival embedded clauses

a. John wants Tom to wash himself.
b. Susan told Mary to respect herself.
c. John asked Tom not to hurt himself.
d. Susan told Mary not to hate herself.

Type C: one-clause sentences

a. John told Tom about himself.
b. Susan talked to Mary about herself.
c. John gave Tom a picture of himself.
d. Susan asked Mary about herself.

Type D: three-clause sentences containing infinitival embedded clauses

a. John thinks that Tom wants Bill to wash himself.
b. Susan believes that Mary told Alice not to hurt herself.
c. John says that Tom told Bill not to hate himself.
d. Susan knows that Mary wants Alice to respect herself.

Vendler Classes and Reinterpretation
Michihiko Kawamura
The Dual Status of the Null Object in Chinese
Yanfeng Qu
On the Representation of Mandarin Syllable Structure
Feng-Lan Kuo
Negative Polarity Items and the Semantics of the Particles
*-to* and *-na* in Korean
Yae-sheik Lee
L2 Acquisition of English Reflexives by Native Speakers of Korean
Hye-Ryun Kim
On the Orientation Problem in Korean *caki* Binding and
the Typology of $X^{\circ}$ Reflexive Binding
Mi-Hui Cho
Xhosa Nominal Tonology: A Domain-Based Approach
Mbulelo Jokweni
Clitics, Case Checking, and Causative Constructions
Xavier Villalba
Complementation of Hausa Aspectual Verbs
Lawàn Danlādì Yàlwa
Deriving the Distribution of Conjunctions
Ed Zoerner

BEST COPY AVAILABLE

# Kansas Working Papers in Linguistics

## Volume 14, Number 2

Studies in Native American Languages

BEST COPY AVAILABLE

*Kansas Working Papers in Linguistics* (*KWPL*) is a regular publication of the Linguistics Graduate Student Association, Department of Linguistics, University of Kansas, Lawrence, Kansas, 66045.

Aim: *Kansas Working Papers in Linguistics* is intended as a forum for the presentation, in print, of the latest original research by the faculty and students of the Department of Linguistics and other related departments at the University of Kansas. Papers contributed by persons not associated with the University of Kansas are also welcomed. The papers published in *KWPL* may not be reproduced without written permission from the Linguistics Graduate Student Association.

Calls for papers usually end in February of each year. Check yearly Call for Papers for exact date. Please send manuscripts, requests for individual orders, and inquiries to:

Editors, *KWPL*
LGSA, Linguistics Department
427 Blake Hall
University of Kansas
Lawrence, Kansas 66045
E-mail: LGSA@kuhub.cc.ukans.edu

The price list for previously published volumes is available upon request. Postage is paid for requests from the United States and Canada. For orders outside the United States and Canada, please add US$3.00 per volume to help defray the costs of postage. A cumulative index may be found in Volume 16 and is available upon request. All institutions with a similar publication are encouraged to request a reciprocal agreement. Standing orders for Number 1, Number 2 (SNAL) or both are welcomed.

The editor would like to express her sincere appreciation to the faculty and staff of the Department of Linguistics for their assistance and encouragement, especially Professor Sara T. Rosen, Mary S. Linn, and Giulia R.M. Oliverio. I would also like to thank the Graduate Student Council for their continuing support.

Cover design, with a minor botanical alteration, by David Andrew Toshraah Nokose Skeeter.

BEST COPY AVAILABLE

# Kansas Working Papers in Linguistics

edited by

Linda M. Roby

## Studies in Native American Languages VIII
Volume 19, Number 2
1994

Partial funding for this journal is provided by the Graduate Student Council
through the Student Activity Fee.

© Linguistics Graduate Student Association
University of Kansas, 1994

# Kansas Working Papers in Linguistics
## Volume 19, Number 2
## 1994

Karankawa Linguistic Materials
Anthony P. Grant....1

Word-Level Nominalization in Choctaw
Marcia Haag....57

Reflexives in Mohawk
Nancy Bonvillain....87

Proto-Algic V: Doublets and their Implications
Paul Proulx....115

The Limit of Structure Preservation in Dak$^h$ota Lexical Phonology
John Kyle....183

Notes on the Kansa Word List of Maximilian, Prince of Wied
Robert L. Rankin....201

# KARANKAWA LINGUISTIC MATERIALS

Anthony P. Grant
University of Bradford, England BD7 1DP

Abstract: In this paper I present the available materials on the diverse dialectal forms of the extinct Karankawa language of coastal Texas in the form of an English-Karankawa vocabulary, together with the attested sentence and text material, a transcription of an alleged Karankawa vocabulary, data on Karankawa personal names, and observations on the use of Karankawa data in Greenberg 1987.

## Introduction.

The purpose of this work is simple and unambitious, namely to make available in compact and usable form the body of known material on the various forms of the speech of the Karankawa Indians of coastal Texas, who have been extinct since the last of their number were slaughtered by Mexican soldiers and Texas Rangers on Padre Island in October 1858. Swanton (1940) published five of the six main sources in a Karankawa-English vocabulary, but his edition shows a number of misreadings, and he omitted the most extensive vocabulary, that collected by Rafael Chowell in 1828-1829, whose existence had already been known for twenty-five years, and which Landar published (with some errors) in 1968.

It should be pointed out at once that our records of Karankawa show the language to have exhibited a very high degree of internal diversity, although I feel that there are enough reconsitutable forms (maybe forty-five) to allow one to regard the data as belonging to forms of speech which were originally a single language.

As far as is possible, I have worked with transcriptions or reproductions of the original vocabularies themselves (in the case of the Talon and Chowell vocabularies), or with early copies (for the Béranger vocabulary), while I have relied on the printed version of Gatschet's materials, which he presumably saw through the press. In other words, this work contains data which are as authentic and as accurate as I was able to make them. I have preserved the original transcriptions of the Karankawa forms and of their French or Spanish glosses throughout, but I have also attempted to provide a (pseudo-)

*Kansas Working Papers in Linguistics*, Volume 19, Number 2, 1994, pp. 1-56

phonemic transcription of the Karankawa forms, leaning heavily on the clues about pronunciation furnished by the spelling of the Karankawa words, especially in the original French and Spanish sources.

This work is based on my 1991 University of Bradford MPhil, 'Karankawa: An Indian Language of Coastal Texas', to which the interested reader is referred for further information on the sociolinguistic background, extensive dialectal variations within the corpus, and philological matters touching upon Karankawa. I have felt it best topresent the forms in an English-Karankawa vocabulary because of problems in disambiguating (and thus alphabetizing!) manuscript forms of certain words in the Chowell, Talon and Béranger vocabularies. My conscious model has been the treatment of Esselen by Shaul, Turner and Collins (1984). I have also included the Karankawa phrases and sentences (mostly from Alice Williams Oliver's memories of apparently xenolectal Karankawa). I have also provided notes on Karankawa personal names, the purportedly Karankawa vocabulary of Guy M. Bryan, and the use of Karankawa data in Joseph H. Greenberg's 'Language in the Americas' (Stanford UP, 1987). The paper ends with a bibliography of works referred to in the text and other relevant works on the subject.

For further information on Karankawa culture, the reader is referred to Gatschet (1891), Schaedel (1949), Newcomb (1961), Wolff (1969), and to Salinas (1990) for information about South Texan tribes generally.

I would like to thank John Gree, Robert Rankin, John Koontz, LeRoy Johnson, David Costa, Carl Masthay, Ives Goddard, Rudolph Troike and John E. McLaughlin for their contribution to this work. None of them are responsible for any errors within it.

## Phonetic Key

This serves two purposes:

1) as a guide to the transcriptional practice of the bracketed forms in the vocabulary

2) as a conversion table in respect to the alphabet used in transcribing Gatschet's material throughout this work. Gatschet used a modification of the Bureau of [American] Ethnology alphabet as laid down in Powell 1880. He did make

some concessions to English usage in Gatschet 1891, such as the employment of <sh ch dsh> for Powell's <c tc dj>, which in American Phonemic are /š č ǰ/.

The following transcription system is biunique. Use of this key will enable the reader to convert from my admittedly ad hoc spelling to the original transcription used by Gatschet. I have in no way adjusted or modified the forms used by Gatschet, even though I think, for example, that his use of doubled consonants is a carry-over from English orthographic usage (the practice occurs in other of his fieldnotes, such as those for Molala), which has no counterpart in reality and that on the other evidence that we have, we cannot be sure that Karankawa had geminate consonants. I also suspect that he overspecified vocalic detail at least in respect to length, though since the rest of the Karankawa material was recorded by francophone or hispanophone observers, whose languages lack phonemic vowel length, we cannot be sure on this point.

I have not bracketed Gatschet's phonetic forms thus [ ], as I have done with the other materials. The reason is simple: the other forms represent my attempt to divine what the recorders thought they heard, in the light of their orthographic usage and the contemporary phonology of their matrix languages (these are not reconstructed or reconstituted forms: it may be that the recorders were wildly inaccurate in their assignation of glosses, interpretation of sounds, and so on). Thus I have taken into account such factors as the greater amount of retention of final consonants, and the differing realisation of *-ille* in seventeenth and eighteenth-century French, rather than simply relying on modern-day pronunciation.

With these other forms I am attempting to guess how a contemporary reader of the vocabulary might have pronounced these words, bearing in mind that the recorders used their native French or Spanish orthographies with few modifications, and left so few overt clues to their transcriptional practices, and that a number of their forms are in any case difficult to decipher and the copies which we have may represent the original field-recordings. In the case of Gatschet's materials, we know what he thought he heard, because he had a semi-phonetic alphabet with which to record his auditory impressions.

It should be emphasised that the only elements of Karankawa which are even vaguely reliable from a phonological standpoint are those which have been reconstituted on the basis of corroboration from more than one source, and which are marked as: R /lams/. For all Gatschet's precision and the attempts of his main consultant at accurate pronunciation, he was not working with native speakers, and phonological cruces should not stand or fall on his evidence alone.

## Consonants

Most of these are as in the IPA system, thus [p] is /p/. The symbol [x] represents a voiceless velar fricative, the so-called German Ach-Laut. In the following remarks, 'G' refers to Gatschet's usage in Gatschet 1891.

| | | | | |
|---|---|---|---|---|
| [j] | G: | <y>: | /j/ | (palatal continuant) |
| [M] | G: | <m̥>: | /m̥/ | (voiceless /m/) |
| [x̂] | G: | <x>: | /ç/ | palatovelar fricative |
| [x] | G: | <x>: | /x/ | velar fricative |
| ['] | G: | <->: | /ʔ/ | glottal stop. |

(<'> is also used in Gatschet: O to represent absence of vowel in [n'] 'I' and [m'] 'you', that is, /n#/ and /m#/: <n' čá áwa> = /n ča áwa/).

[H] is an archigrapheme standing for the unknown and putative value of written <h> in the French and Spanish sources: either /h/ /'/ or /Ø/. Similarly, /VN/ in the francophone sources can be read as /Ṽ/ or /Vn/.

In the phoneticised forms of Spanish recordings, the voiced obstruents have been written as stops: <b d g> for the sake of convenience, rather than as voiced fricatives. Chowell's starred <ch>, here <ĉh>, is /š/; his <ĝn> may be /ŋ/.

## Vowels

The five vowels [i e a o u] have the values assigned to the in IPA. Length is denoted by doubling [ii ee aa oo uu]. [ë] represents schwa. In Gatschet's work [ä] represents the open-mid centralised vowel, written here as /ă/, while [å], occurring ony in the word for <u>crane</u>, represents a low-mid back vowel. [î] occurs in a couple of words in Gatschet's records; he does not explain it, but it was evidently /ii/ (David Costa, personal communication, November 1991).

Because I have used ['] for the glottal stop, in order to ensure the biuniqueness of this representation of Gatschet's spelling, I have employed the symbol [E] between consonants where Gatschet writes <'>, that is, zero, or maybe subphonemic /ë/, thus [glosEn] = G: <glos'n>. Nasal vowels in the transcription are [ã ẽ õ ũ œ̃].

## Stress

Stress is marked with [´] on the vowel, or first vowel of a group except where evidence suggests that the second vowel is to be stressed. I have not marked stress on phonemicised forms of French words, since it is not possible to reconstruct this. In transcriptions of Spanish forms, I have followed Spanish stress rules. Stress in Gatschet's material is irregular, there being no one syllable which customarily receives stress, and thus is probably the best source of evidence about stress placement in Karankawa.

I have preserved the original orthography of the French and Spanish glosses of the Karankawa forms, even in instances where it diverges from modern usage, such as Spanish <ayre> for 'aire', French <cest adire> for 'c'est à dire'.

## Provisional Phonology of Karankawa

It should be pointed out that this is extremely tentative.

The provisional phonemic inventory of Karankawa might be posited as: /p t ts č k kw b d g s š x kl gl l w j m n h '/ for consonants, /i e a o u ii ee aa oo uu/ for vowels, and free stress, with a syllabic canon of (C) V(:) (C) (C). Problematic issues include the nature of /kl gl/ (possibly voiceless and voiced laterals affricates [ƚ λ], although no sources records them as such), the relationship between /ts s š č/, the question of the nature of the sounds reconstructed as /b d g/ (a voiced-voiceless distinction in stops is not an areal feature of this part of Texas, although Coahuilteco and Tonkawa have a voiceless glottalized/unglottalized distincton), the status of other phones, such as /ŋ / (hinted at by Chowell) or /J̆ M/ (the latter = voiceless /m/; these attested by Gatschet), the nature of the sound underlying the <j> of French transcriptions (here transcribed as /ž/), and whether or not /ë/ was phonemic or an allophonic reduced form of one or more vowels.

## Some Remarks on Morphosyntax.

All we have on which to base a model of Karankawa morphosyntax are a few sentences, a couple of short translated texts rendered into a xenolectal and essentially flexionless Karankawa, which was certainly not the way in which the Karankawas themselves spoke, and the evidence of some compound words. Essentially we know next to nothing about Karankawa inflectional and derivational morphology and can only conjecture about its syntax.

To judge from these fragments, Karankawa was an SOV language, with secondary SVO word-order (about 2/3 versus 1/3 of the sentences being SOV). The verb in a subordinate clause always came last (for example in Mrs Oliver's version of 'I want to shoot deer', which is literally 'I want deer shoot').

Affixes (such as the feminine marker /-nen/) seem usually to have followed the noun (see Coyote 1, for example), though otherwise the general order seems to have been Modifier + Head (for instance in Chowell's word for HAT, vide supra). To judge from Mrs Oliver's material, the negative particle preceded the verb.

There may have been pronominal subject or object prefixes on the verb (see Chowell's version of 'I don't want you', in the sentences), and there may have been personal prefixes on nouns (a number of which are attested with forms of a prefix /emi-/), but we cannot be sure of the correct interpretation and assignment of meaning to these forms. Similarly, a few nouns are attested in two or more forms, one of which is often characterised by the affix /-n/, which is also absent in other recordings of the same stem, but we do not know what - if any - significance this affix had, nor do we known anything about possible noun pluralisation. Furthermore, we have no data on which to base observations about whether Karankawa was prepositional or postpositional.

To sum up, Karankawa was a verb-final language, which does not appear to have been, in respect to structure, particularly divergent from other languages in the area.

## English-Karankawa Vocabulary: introduction

This vocabulary, arranged according to English glosses, contains the whole of the lexical material. Its purpose is historical and comparative: it provides etymologies for the minority of items whose origins can be discovered, and through the use of cross-referencing it compares different reflexes for the same gloss in different vocabularies, as well as divergent meanings of the same stem as attested in different sources.

The structure of an entry is explained below.

The lemma or gloss under which an entry is indexed is in CAPITALS; since there is a considerable amount of cross-referencing in this vocabulary, this had been done to make things easier for the reader. If the same gloss has more than one reflex, these are distinguished thus: HOUSE.1; HOUSE. 2.

After the gloss comes the reconstituted form, when such is available, thus: R /teš/. Following this are citation forms from the sources, enclosed in <acute brackets>; the sources recorded by French- and Spanish-speakers are also phonologically explicated in [square brackets], thus: T <tech> [teš]. Spanish-matrix forms from Chowell also have stress indicated according to Spanish stress rules. The phonetic key is explained above in the Introduction. Letters that are difficult to read in the manuscripts and which may have been misread are indicated by underlines.

The symbols used to denote the attestation of the forms are:

T: Vocabulary of the brothers Jean-Baptiste and Pierre Talon, dictated to M. de Boissieu at Morlaix in Brittany in 1689, comprising words learned while the Talon brothers had been captives of the Karankawas around 1686, and originally glossed in French; this is taken from Villers du Terrage and Rivet (1929) and informed by the readings in Troike (1987);;

B: Jean Béranger's vocabulary, 1720-1721, collected near Matagorda Bay, glossed in French; this is taken from Villiers du Terrage and Rivet (1919);

C: Vocabulary most probably collected by the Mexican geologist Rafael Chowell, 1828-1829, somewhere in southern Texas or possibly at Laredo in the state of Coahuila, Mexico, glossed in Spanish; this is taken from Landar 1968, which has been crosschecked with a xerox from microfilm of the original vocabulary in Berlandier and Chowell 1828-1829;

S: Material gathered from Old Simon, a Tonkawa, in September 1884 at Fort Griffin, Shackleford County, Texas, by Albert Samuel Gatschet, glossed in English; this was first published in German in Gatschet (1886), and the present reading is taken from Gatschet (1891);

W: Material gathered by Gatschet from the blind Tonkawa woman, Sallie Washington, also in September 1884 (she had once lived with a Karankawa man); this material is glossed in English; the sources for this are the same as those for the previous item;

O: Material gathered from Mrs Alice Williams Oliver at Lynn, Massachusetts, in 1888-89 by A. S. Gatschet, glossed in English; also from Gatschet (1891), with dual pagination, here preserved, first the page-number of the work as a separate publication, then the (higher) page-number of the monograph as part of a volume.

Other forms are identified by their collector's or source name spelt out, eg, Grasmeyer, or Schaedel 1949, Gursky 1963.

An attempt is made to deal with all attestations of one stem under one lemma. Consequently, there are some cross-references, thus Blanket refers one to Clothing, 2, where C's gloss 'frasada' (poncho) for <lams> is discused with other attestations of the stem in its senses of 'sail, clothings, trousers'.

## English-Karankawa Vocabulary.

ABOUT TO, GOING TO, INTEND TO: G; <čápn>. Gatschet (1891: 78/ 142) claims its derivation from <čá> 'to see'.

ACORN: B; <calache> [kalaš] 'gland de chêne'

AFTER A WHILE, AT PRESENT, SOON: G:<messús, mësus, mEsús>

AIR: C: <lun> [lun] 'ayre'

ALLIGATOR. 1: S: <hókso>

ALLIGATOR. 2: C: <oñase> [onjáse] 'cayman'. Note that both forms resemble Timucua /honoso/ 'crocodile', which Crawford 1988 connects with the widespread Southeastern loanword for 'buffalo', /jánasa/ etc (see Taylor 1976 for further details). It is possible that the two Karankawa words are related to each other and thence related to /honóso/ or a similar form, though this is only speculation.

ALWAYS, ALL THE TIME: O: <mušawáta>

AND. 1: O: <a>

AND. 2: O: <ténno> Note: both these link nominal phrases, see the sentences for details.

ANGRY, I AM VERY ANGRY: S: <napé-nai naxerúaxa pára>, where <napé> = 'I'

ANKLE: C: <iclea> [ikléa] 'tobillo'; cf. HEEL

ARM. 1: B: <se-imahaha> [së-imaHaHa] 'bras' (specifically upper arm?)

ARM. 2: C: <laaje> [la'áxe] 'todo el brazo' (the entire arm)

ARM. 3: C: <ĉhig-mia> [šíg-mia] 'brazo'

FOREARM: B: <se-chotan> [së-šotã] 'bras du coude à l'épaule' (arm from elbow to shoulder)

UPPER ARM: C: <ĉhajegual>, maybe <ĉhasegual> [šaxewál], [šasewál] 'la parte superior del brazo'.

ARROW R /demóo - demúu/: T: <demo> [dëmo] 'flesche' (sic); B: <quechila-demoux> 'balle de mousquet' (ball for musket) [këšila-dëmu] = METAL + ARROW; O: <dëmóo'a>

ASHES: B: <ahona> [aHona] 'cendre'

AXE: B: <quialn> [kian] 'hache' (axe), also 'pickaxe'

BACK OF HAND: C: <cuama> [kwáma] 'la parte superior de la mano'

BAD: O: <čúta> 'bad, dangerous, obnoxious' <áM čúta> is 'octopus'.

BARREL: O: <búudel>. Loanword from Spanish barril.

BASIN: B: <coje eun> [kožœ̃, ? kožõ] 'gamelle'

BEADS (GLASS): B: <qujahin> [kũžaHẽ], or probably <quiahin> [kiaHẽ] 'verroterie'

BEANS, PEAS: B: <coudeche> [kudeš] 'fèves, pois'

A BEAR: O: <ŏs, óoss> < Spanish oso (Gatschet writes the first form with a breve over the <o>, the only instance in his materials.)

BELL: C: <selabaya> [selabája] 'campana'. ? Contains 'metal' root /sel - šel/.

BELLY. 1: B: <a-louc> [a-luk] 'ventre, estomac'

BELLY. 2: C: <coog> [ko'óg] 'varriga' (sic) (belly, bowels). John E. McLaughlin points out the similarity in meaning and shape of Comanche /koe/.

BIG, GREAT, TALL. 1 R /konã - kunẽ/; T <counin> [kunẽ] 'grand'; B: <caa conam> [kaa konã] 'barrique' (cask)
BIG. 2: O: <ja'án>

BIRD. 1 R /kóots..., kúut..., kúud.../; T <tecotsen> [tëkotsẽ] 'un autre oyseau ord're (ordinaire)', listed after <balséhé> 'eagle' (q.v.); cf. B: <coutsen> [kutsẽ] 'alouette' (lark); O: <kódn, kútn. kúdn> 'bird',also 'hen'. Note the Louisiana Siouan language Biloxi /kúudësk/ "bird" (Dorsey and Swanton 1912: 216), although several tribes separated the Biloxis and the Karankawas.

BIRD. 2: C: <coočho> [ko'óšo] 'pájaro', cf. C: <cučhhí> [kuš-Hí] 'pies de pájaro' (bird's feet), <coočhcam> [kó'oškám] 'pluma' (feather). Perhaps related to BIRD. 1, as an example of dialectal divergence. But note T: <cocho> [kóšo] 'calf': could the meaning of BIRD. 2 be 'young of animal, fledgling'?

BISON. 1 R /teš..., tes..., tets.../; T: <tech> [teš] 'boeuf'. <tech-nen> [teš-nẽ] 'vache' (cow); O: <tets'oa> 'cow, cattle, beef' (the only word containing a glottalised continuant recorded for Karankawa) - this last may represent a phonological dialectal isogloss. The stem occurs in C: <tečh-lo-disa> [teš-lo-dísa] 'pig', but a form <tečh> for 'bison' is not attested in C.

BISON. 2: B: <didotte> [didot] 'boeuf'

BISON. 3 R /tik..., tek.../: S: <tikëmai>. cf. B: <teque-dolan> [tekë-dolã] 'cvorne de boeuf' (animal horn). Unlikely to be connected with BISON. 1.

BISON. 4: C: <d̲ola-lá> [dola-!á] 'cibolo', <d̲ola-jay> [dola-xaj] 'cuero de cibolo' (bison hide)

BLACK. 1: S: <má>

BLACK. 2: O: <pal>. Speaker was unsure of this form.

BLANKET: see CLOTHING. 2.

BLOOD: C: <ṯeĉhandelman> [tešandélman] 'sangre'; cf C: <ṯeĉhi> 'flesh'.

BLUE: O: <tsóol>. Perhaps borrowed from Nahuatl *texotli* /tešo'tli/ 'ditto'? Possibly from Spanish *azul* - note Comecrudo /pajasul/ 'blue' from that source.

BOARD: B: <couaham> [kwaHã] 'planche'. Perhaps related to T: <cohal> 'bois'?

BOAT. 1 /..waa../: B: <ouahahim> [? waHẽ] 'pirogue' (dugout canoe); O: <awáan> 'boat'. John McLaughlin points out Numic resemblants in Comanche /waaH/ 'cedar', and Panamint /waa-/ 'piñon pine'.

BOAT. 2: B: <elouchoun> [?? elušũ] 'navire' (ship)

BOAT. 3: C: <ṯualagle> [tualágle] 'buque'. Widespread South Texas loanword? Compare Comecrudo /(pakwa)tataple/ 'canoe'.

BODY: C: <quismata*mac*> <? -*mai*> [kismatamák, -máj] 'todo el cuerpo'

BONE: C: <ṯeĉhedall> [tešedalj] 'hueso', cf. C: <ṯeĉhi> 'flesh', and THIGH.

BOW. 1: T: <crouin> [kruẽ] 'arc', maybe <crouine> [kruin]. According to LeRoy Johnson this is a loan from Cotoname <karua> 'arrow', also recorded in the unpublished collection of South Texan vocabularies assembled by Berlandier and Chowell before 1830.

BOW. 2: O: <gá'i>. Maybe also B: <ca ay> in PISTOL, though this may be the word for STONE, ROCK, or another root. It is just possible that T: <crouin> is an attempt to represent a word beginning with a velar fricative, such as Comecrudo /xái/ 'wood', /xái pataple/ 'bow', in which case it could be a diffused word in southern Texas.

BOWL R /koko - koku.../; B: <coq> [kok] 'écuelle, seau'. Cf. FIREPOT, KETTLE. BOY. 1 R /kloohs - gloos/; T:

<colohs> [kolohs] 'garçon', C: <clox> [klo(h)s] 'muchacho'; O: <glosEn, glo-ëssën>

BOY. 2 also YOUNGSTER: S: <úši níktam> 'little man'. According to LeRoy Johnson (personal communication) this is Tonkawa, but I have been unable to segment it.

BRANDY, WINE, ALCOHOL: B: <clebeu> for <cleben> [klebẽ] 'vin'; C: <liban> [líban] 'aguardiente'; O: <labá-i> 'whiskey' . This word for alcohol is a loan from Lipan Apache, where it is a stative verb meaning 'it is grey', referring to the colour of the local intoxicant tezguino or tiswin, which the Apaches used to sell to other Indians in Texas, and this is cognate with other Apachean and Athapaskan words, such as Navajo /txó-łbáí/ 'tiswin' (water+grey). (See Goddard 1979: 383).

A BRAVE, A WARRIOR R /tešV.../: T: <techoyou> [tešweju. tešoju] 'homme'; C: <ʒečhigua> [teší(g)wa] 'valiente'. Probably derived from the word for FLESH; the connection with BISON. 1, if any,is uncertain (totemism?). It provides evidence that the initial letter in Chowell's handwriting which resembles a modern English cursive J is to be read as <T>. Landar's English gloss 'brave' is a term which is offensive to American Indians.

BREAD. 1: C: <cuama-maya> [kwama-maja] 'el pan', <cuam-pá> [kwampá] 'tortilla'; O: <kwiamója> 'bread'. The words are derived from the root for MAIZE; meaning of <-maja> and <-pá> unknown.

BREAD. 2: B: <cocam> [kokã] 'pain frais' (fresh bread). Presumably related to the word for BOWL

BREAK, TO: O: <táhama>

BREAST. 1: O: <kanín, kenín, kënín>; O: <ka/ënín-ma> = MOTHER. This word for 'breast' is a widely-diffused loanword, originating in Cotoname and spreading also to Comecrudo and Karankawa.

BREAST. 2: C: <al> [al] 'pecho'

BREAST. 3: C: <evem> [ebém] 'los pechos de la mujer'. It is tempting to see this in the manuscript as an emendation for a form originally resembling BREAST. 1, but this is not the case. The original manuscript <E> has had a curve added to make it resemble <C>; the <v> is just a short rightward stroke.

BROAD-FACED MAN: C: <velóo-dulm> [beló'o dulm] 'el hombre frenton'. The <lóo> portion is 'nose'; <dulm> resembles the first part of <dalmac-cama> 'hat' and <daal> 'head'.

BRUSH: O: <tesselénja, -lénia>. Is <tes-> connected with BISON. 1, relating to bison hair?

BUCKET: S: <koláme> 'tin bucket'.From Nahuatl comalli /komalli/ 'frying-pan'?

BUFFALO see BISON

BULL: C: <chool-la> [čo'ól-la] 'toro', also <chool-nen> 'vaca' (cow) and <chool-cuain> 'becerro' (calf)
The forms under BISON. 1 (and BISON. 2 seem also to have referred to domesticated bovines, as forms such as that for MILK would suggest.

BUTTER: C: <ţecha> [téča] 'mantequilla'
Another word connected with BISON. 1? Compare MILK.

BUTTOCK. 1: C: <mooj> [mó'ox] 'nalga'

BUTTOCK. 2 (also PODEX): C: <hačha> [Háša] 'culo'

CALF: T: <cocho> [košo] 'un veau'. According to John McLaughlin, this may be from Comanche /kuhcu/, Shoshone /kuittšu/ 'bison'. For another word for CALF, see BULL. For a possible interpretation of <cocho>, see BIRD. 2.

CALICO: O: <kádla>; <kwíss kádla> 'calico dress' , see CLOTHING. !

CALUMET, PEACE PIPE: B: <cadiolle> [kadjól] 'calumet'.

CANNON: B: <etjam> [etžã] 'canon'

CASK: B: <caa conam> [kaa konã] 'barrique'. <conam> = BIG?

CAT R /gátV/; O: <gáta>, also <gáta kwán> 'kitten'; C: <catum> [katúm] 'gato', <catum-nen> [katũm-nen] 'una gata'. Evidently a loan from Spanish gato, gata. Mobilian Jargon and Choctaw have /katos/ also from Spanish - could this explain the voiceless initial of C's form? It is possible that it is an English loanword (the Muskogean-speaking Alabama Indians now at Livingston, Texas, just north of former Karankawa territory, use /kati/, from English). If so, then proper names apart, it is the only recorded English loan into Karankawa. As a loan, the reconstituted form has no historical depth or great validity.

CATCH, CAPTURE: O: <háitn>

CHEEK: C: <agui> [ági] 'carillo'

CHICKEN: R /ku(d)nVwVla/: C: <connuanguila> [konnuangíla, konjuangíla]/ 'gallo' (cockerel); <connuagnen 'gallina' (hen); O: <kútnewólja> 'hen, prairie chicken'. Gatschet translates O's form as 'BIRD + PLENTY + THERE', but C's recording rather gives the lie to this interpretation. Cognate with and presumably a derivation from BIRD. 1

CHIEF. 1: Schaedel 1949: <Tama> [táma] recorded by Fray Juan Morfi, Fray Gaspar de Solis or possibly Athanase de Mézières (Schaedel's account is unclear).

CHIEF. 2: O: <hálba>

CHILD see BOY

CHIN R /an(j)éna/: B: <em-imian hanéna> [em-imiã (h)anéna] 'menton'; C: <agnena> [agnéna] /? anjéna/ 'barba' (chin, beard). The first form contains a prefix frequently attested in B, but of uncertain meaning.

CHURCH: C: <catssé> [katsé] 'iglesia'. A puzzling form, with a post-Columbian meaning, but with no resemblance to Spanish or Latin forms, or to teopan 'God-house', the Nahuatl form employed in several Rio Grande languages, including Coahuilteco. Hardly to be derived from Latin ecclesia or Spanish iglesia. Franciscan missionaries among the Karankawas had little success, and it is odd to find such a form employed by the Karankawas.

CIGAR: C: <caje-tible> [káxe-tíble] 'cigarro'. See TOBACCO. 1 for first part.

CIGARETTE: S: <ka swénas>. Perhaps 'rolled tobacco'? The Tonkawa word for 'cigarette' is /nepaxkan-pilil/, the first element being 'tobacco' and the second meaning 'to roll' (Hoijer 1949: 24). See TOBACCO. 1 for first part of the

Karankawa word.

CIRCULAR, DISK-SHAPED, ROUND: O: <lá'akum>

CLAVICLE, COLLAR BONE: C: <lem-ḏolmá> [lem-dolmá] 'clabicula' (sic). Connected with HEAD. 2

CLOTHING. 1: O: <kwíss> 'dress'; <kwíss kádla> 'calico dress'; also <gus-> in O: <gusgáma> 'shirt'
A loanword from Comanche /kwásu'U/ (Robinson and Armagost 1990), which is common Numic and is recorded in Southern Paiute, Bannock and other Numic languages of the Great Basin. The item and name were presumably borrowed through trade.

CLOTHING. 2 R /lams/; T: <alames> [alamés] 'l'habis d'un european' (sic); B: <a-lams> [a-lams/-lãs] 'vêtement' (clothing), <em-lams> [em-lams /lãs] 'voile' (sail); C: <lams> [lams] 'fresada' (poncho, blanket). Cf. also C: <lams-santle> [lams-sántle] 'pañuelo' (handkerchief).

CLOTHING. 3 R /...kama, ...gama/; C: <cama> [kama] in: <dalmac-cama> [dalmák-káma] 'sombrero' (hat) (< head-clothing); <ţenno-cama> [ténno-káma] 'pantalon' (trousers) (? hip-cloth), <ĉhacama> [šakáma] 'camisa' (shirt). The word for SHOE is evidently derived from the same stem.

CLOUDS: C: <quapan> [kwápan] 'nubes'

COAL, LIVE: C: <alm> [alm] 'braza' (for 'brasa').

COLD R /glaj - glẽ/: B: <delin> [dẽlẽ] 'froid'; C: <gláy> [glaj] 'frío'. Swanton's phonemicised reading of [delin] for B masks what is possibly evidence for a voiced lateral fricative in Karankawa.

COME. 1 R /gá'as/: O: <gás>; W: <ká'as wána> 'come here!'

COME. 2 R /ewé'e/: W: <ewé'e> 'come! come quick!', S: <éwee> 'come!, come here!' Probably an interjection rather than a true verb.

CORD: B: <bachina> [bašina] 'corde'

COWARD: C: <ţeĉhi-ĉhi-salem> [tešišisalém] 'cobarde'; cf. FLESH,; maybe <salem> is connected with C: <saylá> 'man'.

COYOTE: C: <cuba> [kúba] 'coyote', also <cuba-nen> 'coyota'. Possibly a loanword from Cotoname or Comecrudo, cf. Cotoname <couba-ajâ> 'lion', Comecrudo <couepet> 'mountain lion' (see Goddard 1979: 378).

CRANE. 1: B: <coln> [koln ?] 'grue'
CRANE. 2: O: <këdóod>; <këdâd>. Possibly related to above? The only word which contains /â/. See Comecrudo /kol/ and Cotoname /karakór/,

CURLEW: B: <quojol> [kwožol - kwežol] 'bécasse de mer'. Cf. DUCK.

DART see METAL

DAY: O: <bákta>

DEAD: O: <mál>

DEAR: O: <mutá>

DEER R /..dotsVn, ...dočin/; T: <tecomandotsen> [tëkomãdotsẽ] 'un chef ou chevreuil' ; C: <esdochin> [esdóčin] 'venado', <esdochin-nen> 'venada'; O: <dó'atn - dó'ëtn - dóotn> (also in NINE and TEN). T contains BISON. 3; variation between <t ts ch> in forms could be indicative of an isogloss.

DEERSKIN: B: <quesoul> [kësul] 'peau de chevreuil'. Cf. GUN, WOOD.

DO: O: <kosáta>; also 'to build' (a fire)
DOG R /keš - kes/; T: <quez> [kez] 'chien', also 'loup' (wolf); B: <queche> [keš] 'chien'; C: <qüečhe; qüečhe-nen> [sic; kwéše, kwéše-nen] 'perro, perra'; O: <kíss>. Also T: <quez calbasses> [kez kalbasës] 'cochon, cest a dire, chiens des Français' (pig; dog of the French). According to Gatschet (1891: 44/ 108), the Karankawa dog resembled the fox or coyote. Perhaps the word is an early loan from Chitimacha /kiš/ 'dog' ; note also Cotoname <kissa> 'fox'..

DON'T CRY, HUSH (exhortation to children): O: <ăhămmiš>

DRINK R /kwẽ - kwe/: B: <coouaen> [kuwẽ] 'boire'; O: <akwetén>

DUCK: B: <coué> [kwe] 'canard'

CANVASBACK DUCK: O: <mëdá'u - medáu - mëdáu>

EAGLE: T: <balséhé> [balseHe] 'aigle'

EAR: C: <aig> [aig, aij] 'oreja'; also <aisoyna> [aisójna] 'oido' (sense of hearing), <aigenal> [aixenál] 'cerilla del oido' (earwax).

EARTH: ? T: <caham-quéamy> [kaHã-keami] 'les espagnols, comme a dire, - gens de terre, parce quils sont allez a eux par terre'' ('the Spaniards - people of the earth, so called because they came overland while the French reached Karankawa territory by sea'). People have usually taken the <caham> portion to mean 'earth', cf. B: <cohon> 'sable - sand'. There is no positive proof for the exact meaning of this phrase and no parallels by which we can establish the meaning of the elements.

TO EAT: O: <aknámas, aknámus>

EGG: O: <dáhome>; Gatschet links this with O: <dă> 'oyster'.

EIGHT: O: <haíkia béhema>; the two elements mean respectively TWO and FATHER.

ELBOW: C: <deeya> [de'éja] 'codo'

EYE. 1 R /le..., la.../: B: <im-lahoué> [im-lawe] 'sourcils' (eyebrows). C: <leca> [léka] 'ojo'. Cf. C: <? leguemb> [legémb] 'pestaña' (eyelash), <lecuim> [lekwim] 'pupila del ojo' (pupil of eye), <lesayma> [lesájma] 'parpado' (eyelid), <lesáyom> [lesájom] 'ceja' (eyebrow).
EYE. 2: B: <emicous> [emikus - emi-kus? em-ikus?] 'oeil'

FACE : C: <ʈancú> [tankú] 'cara'

TO FALL: O: <amóak>

FAR OFF: O: <njá wál> = THERE + MUCH

FATHER: O: <béhëma - béhema - béx̂ma>. Derived from PENIS; this word for FATHER may be a replacement for a tabooed word. Cf. also MOTHER. Also occurs in the numeral system at FIVE and EIGHT - perhaps a misremembering for <étsma> 'hand, finger'.

FEATHER. 1: B: <hamdolocq> [Hãdolók] 'plumes'. Maybe includes a personal prefix on the root /dol/ also attested in TOOTH, HORN. Feathers, especially those from eagles, were trade and prestige items in the Plains culture area.
FEATHER. 2 - see BIRD. 2

FEMALE R /...nen/: T: <-nen> [-nẽ] on BISON. 1, C: <-nen, -nem> [? -nen, -nem] on DOG, CAT, CHICKEN, TURKEY, BULL, WOLF, COYOTE, HORSE. Perhaps a bound morpheme? Note Cotoname /nan/ 'female'.

FIELD: C: <dopsa> [dópsa] 'prado'

FINGER : O: <étsma>, also 'hand'

FIRE. 1 R /kwejle/: T: <cohoille> [kwojlë, kwejlë] 'feu', B: <quoylesem> [kwoj/ejlesẽ] 'feu'

FIRE. 2 R ? /kwášV/ (perhaps /kwaša - kwači/: C: <cuačha> [kwáša] 'fuego'; O: <kwáči>. The latter is also a male proper name.

FIRE. 3: S: <húmhe>

FIREPOT: T: <coko> [koko] 'pot a feu'; cf. BOWL, KETTLE, maybe also FRESH BREAD

FIRESTICKS: B: <acta demajé> [akta dëmaže] 'bois avec lequel on obtient le feu par frottement'.

FISH. 1 R /am/; C: <am> [am] 'pez' (live fish); O: <áM>. FISH. 2: B: <quyles> [kils] 'poisson'

FIVE: O: <náatsa béhema> literally ONE + FATHER (? mistake for <náatsa étsma - one + hand)

FLAGON: B: <quedica> [këdika] 'flacon'

FLOUR: O: <ámhătn>; <já ámhătn> = 'cornflour' (<já> = <jáam> 'potato')

A FLY: B: <cameje> [kamež] 'mouche' FOOT. 1 R /...eHV.../ ; B: <eham> [eHã] 'pied'; C: <hei-yú> [Hei-jú] 'pie'; <cučh-hí> [kuš-Hí] 'pies de pájaro' (feet of a bird) , <hei-yosam> [Hei-josám] 'los dedos del pie' (toes); maybe also O: <kékeja> 'foot'. An awkward set to reconstruct - if, indeed, the forms are cognate.

FOOT. 2: C: <del> [del] 'toda la pierna'; cf. THIGH

FOOT. 3: C: <ik - ic> in <ik-dota> [ik-dóta] 'talon' (heel), <iclea> [ikléa] 'tobillo' (ankle) and <ik-aal> [ik-aál] 'planta del pié' (sole of the foot). Possibly a bound morpheme.

FOREHEAD: C: <meklóo> [mekló'o] 'frente'; Cf. NOSE

FOUR: O: <hájo hakn> 'four'; cf. TWO.

FRENCH: T: <calbasses> [kalbasës] 'français, comme qui diroit, gens venus de la Mer' ('French, that is to say, people from the sea'). This word is unanalysable. Nothing to do with Louisiana French çalebasse 'calabash, gourd'! T: <quez calebasses> [kez kalbasës] is 'cochon' (pig) - 'French dog'.

FRIEND. 1: O: <ahájika>. Possibly related to MAN. 2 if reading of latter is <ahax>.

FRIEND. 2: O: <amígo>. Spanish loanword into Karankawa which was replacing FRIEND. 1 in 1840's, according to Mrs Oliver.

FUCK: C: <haĉhi cooĉhe> [Háši ko'óše] ' vamos á joder?' (sic) (are we going to fuck, shall we fuck?)

One of the few genuine Karankawa sentences recorded. <cooĉhe> could be related to O: <kosáta> 'to do'; <haĉhi> is presumably C: <hačha> 'rear'; the metaphor is attested in other languages, such as Vlach and Balkan dialects of European Romani and polisson colloquial English.

GET AWAY!: O: <ăhămmiš snín!>. <snín> is attested only in this phrase.

GIMLET: B: <clany> [klani] 'vrille'

GIRL: O: <káada>; also term of address by mother to daughter.

GIVE R /baHúš/; C: <b̲ajuĉh> 'dime'; O: <báwûs>

GLASS see METAL

GLASSWARE see BEADS

GLOVE: C: <oñecadá> [onjekadá] 'guante'. Cf. ALLIGATOR. 2

GO. 1: O: <jé>. Also see COME.

GO. 2: O: <budáma> 'gone, past'.

GO AWAY: W: <wána> translated as 'go away' when isolated, but as 'come here' when the phrase is <ká'as wána>

GOD: C: <dios> [diós] 'Dios'. Loanword from Spanish. Two names of Karankawa gods according to Morfi (in Schaedel 1949) were Pichini and Mel.

GOLD: C: <ĉhelee-ĉheman> [šelé'ešéman] 'oro'. Contains widespread root for METAL.

GOOD. 1 R /baHa/: T: <couist-baha> [kwist-baHa] 'bon, quelquechose de bon'. B: <baa> [ba'a] 'expression de contentement'. <couist> also occurs in Talon's Caddo list in word for 'good' : <couhistanhat>. The second half parallels modern Caddo /ha'ahat/ 'good'; the first is a mystery. It has no parallels with 'good' in Caddo or Karankawa, and its only resemblance is to /wísts'i/, the modern Caddo word for 'one', which looks like a loanword from the Dhegiha Siouan language Osage, cf. Osage /wíhtsi/ 'one' (other Dhegiha languages, Quapaw, Kansa and Omaha-Ponca, have forms beginning with /m-/). The earlier Caddo word for 'one', */kawanay/, went out of use in the mid-nineteenth century. However, <couhistanhat>, or /ku-wisc'-ha'ahat/, meaning presumably 'something good', does not seem to be a grammatical sequence of morphemes in Caddo.

GOOD. 2: O: <plá> 'good, good-looking, fine, useful'. Possibly connected with Comecrudo /peléx/ 'good'. GOOD BYE, FAREWELL: O: <ačáta>; derived from *[náyi áawa ča ta] 'I + THOU + SEE = WANT = I want to see you' as Gatschet suggests?

GOOSE: O: <lá'ak>. Probably imitative or onomatopoeic. There are similar forms throughout North America from Tunica on the Gulf Coast to Yana and Chimariko in Northern California.

GRASS. 1: C: <quay> [kway] 'zacate'. Rafael Chowell, the collector of the list, states that he could hear no difference between this word and the one for 'horse'.

GRASS. 2: S: <awáčxol>.

GRINDSTONE: B: <hama> [Hama] 'meule'

TO GROW: Derived from SMALL. O: <kwáan, kwánnakwan, kwánakwan>

GUMS OF TEETH: C: <eclenemac> [eklenemák] 'encías'

GUN. 1: B: <quisoulp> [kisul?] 'fusil'; cf. WOOD, DEERSKIN. Probably a misunderstanding on Béranger's part.

GUN. 2: C: <ĉhelacuy> [šelákuj] 'un fusile' (sic). Cf. METAL

GUN. 3 see PISTOL

GUNPOWDER R /ku/onmel/; B: <calmel> [kalmel] 'poudre à fusil', O: <kúunmil>; C: <con-mel> [konmél] 'polvoro'. An unusual amount of agreement about a term used to describe a post-Columbian item. Could [kon] be the same word as <cohon> 'sand' and the whole compound mean 'black sand, black earth'?

GUTS. 1: C: <ɟaĉh> [taš] 'tripas'; <traĉh-sá> [taš-sá] 'azadura' (animal guts, chitterlings). The two terms differ clearly in their initial letters, though the first is not easy to read. One of these is probably a misrecording.

GUTS. 2: C: <coog> [ko'óg] 'varriga'. See BELLY. 2

HAIR R /ekwa - ekwe/: B: <equia aycouy> [ekia ekwi] 'les cheveux'; C: <equa> [ekwa] 'cabello', <ecun-eĉhe> [ékun-éše] 'un solo pelo' (a single hair).

HAND. 1: C: <hooyo> [Ho'ójo] 'mano', cf. <? hooyo-am> [Ho'ójo-am] 'dedos de la mano (fingers), <? hooymblé> [Ho'ójmblé> 'uñas' (fingernails), <ho-yal> [Ho-jál] 'palma de la mano' (palm of hand).

HAND. 2: O: <étsma> - see FINGER

HANDKERCHIEF: C: <lams-santle> [lams-sántle] 'pañuelo'. Second part unidentified; first part is CLOTHING. 2

HANDSOME, BEAUTIFUL, PRETTY R /...málV/; O: <hamála>; cf. C: <calee-malem> 'muger bonita' (pretty woman).

HAT. 1: C: <dalmac-cama> [dalmák-káma] 'sombrero'; <cama> occurs in other names of garments, while the first element is derived from the stem for HEAD.

HAT. 2: B: <calama> [kalama] 'chapeau'. Possibly connected with the above.

HATCHET: O: <mačíta>. Loanword from Spanish machete.

TO HATE: O: <matákia>

HE (also SHE, IT, THIS, THAT): O: <tál, táll>. Probably a demonstrative rather than a true third-person pronoun.

HEAD. 1: B: <enoquea> [enokea], probably [en-okea] 'tête'

HEAD. 2: C: <daal> [da'ál] 'cabeza'; cf. <velóo-dulm> 'broad-faced' and <dalmac-cama> 'hat'.

HEALTHY: O: <klabán>

HEART: O: <láx̂hama - láhama>

HEAT: C: <? sĉhoj> [? šox] 'calor'

HEEL: C: <ik-dota> [ik-dóta] 'talon'; cf. ANKLE

HERE: O: <njá, niá> = THERE

HIDE, SKIN: C: <ḏola-jay> [dóla-xaj] 'cuero del cibolo' (buffalo hide) (<ḏola> = 'buffalo')

TO HIT see TO STRIKE

HOE: C: <ĉhelee-nagut> [? šelé'e-nagút] 'hazadón' (sic); cf. METAL

HORN. 1: B: <tequedolan> [tekëdolã] 'corne de boeuf'. For <teke> see BISON. 3; <dolɵn> also occurs in B: <dolonaquin> 'tooth' and possibly in B: <hamdolocq> 'feathers' - could it be a stem meaning 'protuberance'?

HORN. 2: C: <homo> [Hómo] 'cuerno'. Loanword? Cf. Comecrudo /jëmó/ and Cotoname /jómo/ 'horn'.

HORSE R /kuwáj/: T: <cauouaium> [? kawájũ]; C: <cuay> [kwaj] 'caballo', also <cuay-nen> [kwaj-nen] 'yegua' (mare), <cuaanñam> [kwa'áñam] 'potrillo' (colt),

<cuaflekuen> [kwaflekwén] 'macho' (stallion); O: <kuwáji, kuwái>; S: <kwan, kwá> (kwán péka. 'white horse', <kwá má> 'black horse'; Grasmeyer has <Qwy> [kwaj]. A loanword from Spanish *caballo*, probably via Nahuatl *cahuayo* /kawájo/. T adds a note: 'Ce toutes les nations sauvages generalement l'appelle ainsy'; however, the word is not used in Tonkawa, Caddo or Atakapa (though it does occur in Caddo (Chafe n.d), and in Wichita /kawáarah/ - David Rood p.c.).

HOUSE. 1 R /káHa/: T: <cahum> [kaHœ̃] (Troike 1987: 294), <caham> [kaHã] (Swanton 1940: 126) 'cabane' (hut); C: <caha> [káHa] 'casa'. Rivet's derivation from Numic, cf. modern Comanche /kahnl/, is unlikely.

HOUSE. 2: O: <bá'ak>. Also BUILDING, CABIN. CAMP, HUT

HUNGRY: O: <amél>, emphatic <améeel>

HURRY. 1: O: <kóta, kotá>

HURRY. 2 R /xanké - xankí/; W: <xankéje>, S: <xankí>

TO HURT, INJURE: O: <kassičúwakn> (Gatschet spells it <kassidshúwakn> [kassiǰúwakn], the only instance in the corpus of [ǰ]). Derived from TO POUND, with unidentified suffix.

I, ME, MY, MINE R /na-/: O: <náji, ná'aji, ná-i, ná-î, n'>; S: <napé, napél>

INDIAN: C: <ĉhuyĝnea> [šuyŋéa] 'indio'; <čhuyǧnea-calem> 'india' (Indian woman). Could this be the native tribal designation of the Karankawas?

IRON: C: <čhelneday> [šelnedáj] 'fierro' [sic]. Another word derived from the common root for METAL

JUG: B: <cahan> [kahã] 'cruche' TO JUMP: O: <éM>

KARANKAWA?: Gursky 1963: 29 <carancagua> [karanká(g)wa] 'los que les gustan los perros' (those who like dogs), allegedly itself a Karankawa designation. If the translation is correct, it confirms Gatschet's guess in Gatschet (1891: 43-44/ 107-108) that the name *Karankawa* is from Comecrudo /klam/ 'dog' and /káwa/ 'to love'. Note T: <Clamcoëhs> 'Karankawa Indians'.

KETTLE: B: <couqujol> [kuküžol] 'chaudière', possibly <couquiol> [kukiol] 'chaudière'. Cognate with BOWL and FIREPOT.

TO KILL: O: <ahók - ahúk>

KNEE R /kla(h)s/: B: <en-clas> [en-klas] 'genou'; C: <clax> [klas - klaks] 'rodilla'.

KNIFE. 1: T: <bequecomb> [? bekëkõb - bekkõb] 'couteau'

KNIFE. 2 R /silV - šelV/; B: <cousila> [kusila] 'couteau'; C: <ĉhela> [šéla] 'cuchillo'; O: <silekáji> 'knife'. Connected with the root for METAL, in this case perhaps influenced by Spanish cuchillo 'knife' - metal tools, including knives, were favoured trade items in the area as they were superior in strength and speed to local stone, obsidian and bone technology.

TO KNOW. 1: O: <kwás - kwáss> TO KNOW. 2: S: <kúmna>. Less common than the above word, according to Mrs Oliver.

LARGE-NOSED: C: <lóo-dulm> [ló'o-dulm] 'narizon'; see NOSE. Possibly <dulm> is connected with <daal> 'head'

LARK: B: <coutsen> [kutsẽ] 'alouette'; cf. BIRD

TO LAUGH: O: <káita - katá>

LEG. 1: B: <em-anpocq> [emãpok/ ãpox] 'jambe'

LEG. 2: C: <sĉhemi> [šémi] 'pierna'

LEG. 3: C: <del> [del] 'toda la pierna'. See THIGH

LET US GO! (Allons!): W: <wána> = COME HERE.

TO LIE DOWN: O: <wúak> LIGHT: C: <est-day> [est-daj] 'la luz'.

LIP: C: <aggmaĉh> [agmáš] 'labio'; cf. MOUTH

LITTLE, SMALL, YOUNG OF AN ANIMAL. 1 R /kwan - kwaan/; B: <ca ay couan> [ka aj k(u)wã] 'pistolet' (pistol - small bow); C: <cali-cuan> 'muchacha' (small woman),

<chool-cuain> 'becerro' (calf - small bull). O: <kwán, kwáan> 'small'; see also TO GROW. Gatschet rather unconvincingly derives this word from <káhawan> 'to make'.

LITTLE. 2: S: <níktam> in <úši níktam> 'boy, youngster, little man'

LONG AGO: S: <upáat>, emphatic <upáaat>

TO LOVE. 1: O: <ka>

TO LOVE. 2: C: <qüaĉhel> [kwašél] 'querer' ('to love' or 'to want'), <mi-qüaĉhals> [mi-kwašáls] 'no te quiero'. Possibly cognate with FIRE. 2? Metaphorical origins of terms for 'love' are not unknown in Native American languages; Pitkin 1985 attests to 'love' in Wintu (Northern California) being derived from verb meaning 'to spin, revolve, whir'.

MAIZE, INDIAN CORN R /kwejam - kwajam/; B: <couejam> [kuežã] 'biscuit', O: <kwiáM> 'corn', <kwiamója> 'bread'; C: <cuayam> [kwajám] 'maíz', <cuama-maya> [kwáma-mája] 'pan', <cuam-pá> [kwampá] 'tortilla'.

TO MAKE: O: <káhawan - ká'awan> 'to make, manufacture'. See also TO DO.

MAN. 1: T: <techoyou> [tëšóju], cf. BRAVE

MAN. 2 R /aha... ??/; B: <ahax - alax> [aHa(k)s - ala(k)s] 'homme' (= vir). If <ahax> is correct reading, then perhaps related to O: <ahájika> 'friend'.

MAN. 3: O: <jámawe>. Possibly a loanword from Cotoname /jómo/ 'horn, man' (metaphorical *pars pro toto*!)

MAN. 4: C: <saylá> [sajlá] 'hombre' (= vir).

MAN. 5: S: <úši>, in <úši níktam> 'youngster, little man'. Can all these be words for 'man'? It is possible that some of these are personal names, misinterpreted as words for 'man'.

TO MARRY: O: <mawída>. From Spanish *marido* 'husband'; 'to marry' is *casarse*. MAST: B: <enguesoul> [engësul] 'mât'. Cf. WOOD.

MAT: B: <didaham> [didaHã] 'natte (tapis)'

MEAT: C: <ţeĉhi> [téši] 'carne'. See BISON. 1

METAL R /šel..., sile.../; T: <techheillé> [? tëšéjlje] 'sabre'. B: <cousila> [kusila] 'couteau, dard, fouène, harpon' (knife, dart, harpoon, pick-axe). <cousilam> [kusilã] 'verre' (glass), <cousilea> [kusílea - kusílja] 'herminette' (adze), <quesila-conan> [kësila-konã] 'assiette d'étain' (tin plate); C: <ĉhela> [šéla] 'cuchillo' (knife), <ĉhelacuy> [šelákuj] 'fusile' (sic; rifle), <ĉheledanic/ ĉheledame> [šeledaník / šeledáme] 'plata' (silver), <ĉhelneday> [šelnedáj] 'fierro' [sic] (iron), <ĉhelee-ĉheman ?> [šelé'e-šéman] 'oro' (gold), <ĉhelee-nagut> [šelé'e-nagút] 'hazadón' [sic] (pickaxe). O: <silekáji> 'knife'. The variation in the root /šel - sil/ is probably dialectal. The B forms seem to be influenced by Spanish <u>cuchillo</u> 'knife'. Maybe the second half of the word for TO SEW is connected.

MILK. 1: O: <tesnakwája>. Cf. BISON. 1

MILK. 2: C: <? sĉhimuaim> [šimuáim] 'leche'

MOLASSES: O: <téskaus gllé'i> = SWEET WATER; see SUGAR, WATER

MOON. 1: B: <a-ouil> [a-wil] 'lune'. Perhaps cognate with second half of <ɗó'owal>, O's word for SUN

MOON. 2: C: <tayk> [tajk] 'luna'.

MOSQUITO: O: <gă, gáx̂>

MOTHER: O: <kënínma, kanínma>. Meaning 'with breasts'? cf. BREAST. 1. A noa word, that is, a cover for a taboo term.

MOUNTAIN: C: <? euajadan> [ewaxádan] 'montaña'. This word is difficult to read in the manuscript. Possibly <cuajadan> [kwaxádan] or <enajadan> [enaxádan], or a form ending in <-au> [-aw]?

MOUTH R /akw..., ag.../: <emi aqouy> [emi-akwe] 'bouche'; C: <agg> [ag] 'boca'. Cf. LIP.

MUCH: O: <wól, wóll, wÁl, wál>. Also means MANY, POWERFUL. MUSCLE: C: <eel> [e'él] 'muzlo' [sic]. See SHOULDER. 2?

MUSIC: O: <jóota>

MUSICAL INSTRUMENT?: A woodwind instrument called <avacasele> is mentioned by Morfi (in Schaedel 1949) as being played at sacred dances. If the word has been read correctly in the manuscript, and if it is indeed a Karankawa word, there is some resemblance between <casele> and the word for WOOD.

MUSKET BALL: B: <quechila-demoux> [këšila-dëmu] 'balle de mousquet'. Literally METAL + ARROW. An acculturation term.

NAILS (FINGER): C: <hooymblé> [Ho'ójmble] 'uñas'. See HAND. 2

NAVEL: C: <ay> [aj] 'ombligo'.

NECK R /sebe - sebi.../: B: <em-cebecq> [em-sëbek] 'cou'; C: <sebillool> [sebiljo'ól] 'pezquezo' (for 'pescuezo')

NEEDLE. 1: B: <? besehena> or <beschena> [bëšëHena, bëšéna] 'épingle' (pin). Swanton 1940: 125 reads <beseceba> [beseseba] for the word as written in the manuscript used by Villiers du Terrage and Rivet (1919), Cf. B: <bachénana> 'tatouages' (tattoo marks).

NEEDLE. 2: O: <aguíja>. Loanword from Spanish aguja [agúxa], often [áwha] in Mexican Spanish. But maybe from French aiguille [egüij].

NINE: O: <háikia dóatn>. Literally TWO + DEER. An example of subtractive reckoning? Atakapa has terms meaning 'hands without two' and 'hands without one' for EIGHT and NINE.

NO, NOT R /k(w)om - k(w)oM/; O: <kúM, kóM, kóoM>; S: <kwóom, kwó'om>. Precedes the verb.

NOSE R /lóo - lúu/; B: <em-ay alouacq> [emi-aluak] 'nez'; C: <lóo> [ló'o] 'nariz', note also <meklóo> 'frente', <velóo-dulm> 'frentón'. See also middle segment of PIG. 2.

NOW: O: <ašáhak> 'now, at present'.

OAK APPLE: B: <aixquitoula> [e(k)skitula] 'pomme de chêne'.

ONE: O: <náatsa>, also occurring in FIVE and SEVEN.

OYSTER: O: <dă>

PADDLE (FOR PIROGUE): B: <em-louajem> [em-luažẽ] 'pagaie'.

PALATE: C: <elcon> [élkon] 'paladar'

PALM OF HAND: C: <ho-yal> [Ho-jál] 'palma de la mano'. See HAND. 1.

PAPER: B: <imetes acouam> [imẽtes akwã] 'papier'. <ime> is probably a personal prefix. Is <tes> connected with BISON. 1 in the sense of 'buffalo hide' (Caddo /núšt'uh/ 'paper' meant 'hide, skin' earlier last century, cf. Taylor 1963: 52). Perhaps <acouam> is connected with O: <kwá, gwá> 'to read'.

TO PASS FROM ONE SIDE TO THE OTHER R /lon/: B: <lon> [lõ] 'passer d'une côté à l'autre'; Gursky 1963: 30 <lon> 'pasar de un lado al otro'.

PEACE: C: <biase> [biáse] 'paz'.

PELICAN: B: <aucmane> [okman] 'grand gosier, pélican'

PENIS R /bax/: B: <emibacq> [emibak] 'organes génitaux masculins'. C: <baj> 'miembro biril' (sic). John McLaughlin notes a Numic resemblant in Panamint /pakan/, morphophonemically [paɣan] 'arrow, penis'.

PIG. 1: O: <madóna>

PIG. 2: C: <teĉh-lo-disa> [teš-lo-dísa] 'puerco', cf. BISON. 1, and NOSE (or maybe STOMACH. 2). Possibly 'bison with a broad snout'. SOW: <teĉh-lo-nem> (for <-nen>); not glossed by C..

PIG. 3: T: <quez calbasses> [kez kalbasës] supposed to mean 'dog of the French'.

PIG. 4: W: <tapšewá> 'hog'.

PIMENTO: B: <quesesmaille> [kesesmajlj] 'piment' (the Ayer manuscript has this as reading 'pennache' i.e. pinnace, small boat)

PISTOL: B: <ca ay couan>. [ka'aj kwã] 'pistolet'. Literally 'small bow'.

PLACE ?: This relates to a pseudo-Karankawa form. Gursky 1963: 32 <tampacuas> [tampákwas] 'lugar de los pintos o tatuados', <tam> = 'lugar' (place), which is identified as being Karankawa. This is recorded as the name of a tribe. But Salinas 1990: 63 correctly claims that this form is Comecrudo: see Swanton 1940: 85 pakahuái, pakawái 'to write, tattoo'. However, despite the origin of the term, Salinas (op. cit: 146) claims that the Tampacuas did not speak Comecrudo.

PLAICE: B: <ampaje> [ãpaž] 'plaise'

PLATE: B: <quesila-conan> [kësila-konã] 'assiette d'étain' (big metal?).

PLOVER: B: <cebé> [sëbé] 'pluvier'

POTATO: O: <jáM>. Probably originally referred to a native tuber, but hardly English 'yam' or Spanish ñame. Also in <já ämhätn> purportedly meaning 'cornflour'. John McLaughlin suggests a parallel in Shoshone /jampa/ 'generic tuber'.

TO POUND, CRUSH: O: <kássig>. See also HURT, INJURE.

PRIEST: The eighteenth-century observer Fray Morfi (cited in Schaedel 1949) records the term <coma> [koma] to refer to priests of native religion.

TO PUSH: O: <dán>

RAIN R /wiis/, maybe /wi'is/: C: <güis> 'llovia' [sic]; O: <wíasn>

TO READ: O: <gwá, kwá>. The meaning is certainly post-Columbian. Possibly with original meaning 'to look at'?

RED: O: <tamójika>

RIBS: C: <guen> [gen] 'costillas'.

ROCK, STONE: C: <cay> [kaj] 'piedra'.

TO RUN: O: <tólos, tólus>. See also words for HURRY

SABRE, SWORD: T: <techheillé> [tešélje] 'sabre'. Connected with METAL.

SADDLE: C: <ţoyaiene> [tojajéne] 'silla ó asiento' (seat or saddle).

SAIL: B: <em-lams> [em-lams? ẽ-lãs?] 'voile'. See CLOTH. 2

SALT. 1: B: <quetache> [këtaš] 'sel'

SALT. 2: Cf. <dem> [dẽ] in T: <comcomdem> 'sea, ocean, literally: saltwater'. Note Cotoname /da'ën/ 'salt'. <comcom> may have an Atakapa etymology, but <dem> does not.

SAND: B: <cohon> [koHõ - kwõ] 'sable'. See EARTH, GUNPOWDER.

TO SAY: O: <káupn>.

SEA. 1: T: <comcomdem> [kõkõdẽ] 'la mer. cest a dire Eau salée' (sic); derived from WATER. 1

SEA. 2: C: <ţacui> [tákwi] 'el mar'

SEA. 3 see WATER. 1

TO SEE. 1: O: <čá, čé>. See also ABOUT TO.
TO SEE. 2: C: <om> [om] 'ver'.

SEVEN: O: <háikia náatsa> Contains roots of TWO and ONE.

TO SEW: B: <tecsilea> [teksilea] 'coudre'. Perhaps connected with <tequedolan> 'horn' and maybe the root /sil/ 'metal'?

SHIRT: C: <Ĉhacama> [šakáma] 'camisa'; O: <gusgáma> 'shirt'. Both forms are partly identical, containing /káma/ CLOTHING. 2. O also contains <gua> from <kwíss>, see CLOTHING. 1

SHOE R /kamepVl/: B: <cameplan> [kaměplã] 'soulier'; C: <camepel> [kamepél] 'zapato'. Possibly connected with CLOTHING. 3. Maybe an unidentified loanword; the Karankawa were known to the Tonkawa as

/kapay-yakokxon/ 'the people without moccasins'.

TO SHOOT: O: <óodn - úudn>

SHOULDER. 1: B: <em-sehota> [em-sëHota] 'épaule'. Connected with and probably identical to <sechotan> [sëšotã], or to <sehotam> [sëhotã] 'forearm'.

SHOULDER. 2: C: <eel-em> [e'él-em] 'hombro'.

SICK. 1: B: <a-eas, a-las> - as read by previous scholars, but perhaps <? a-cas> [ajás. alás, akás] 'malade'. The manuscript reading is unclear.

SICK. 2: O: <kwáčo, kwátsu>. Related to <kwáči>? See FIRE. 2

SILVER: C: <ĉheledame / ĉheledanic>. [šeledáme / šeledaník] 'plata'. See METAL.

TO SIT R /hákV/; O: <hákus, hákës>; W: <háka>.

TO SIT DOWN: W: <čakwamé> 'sit down!'. Perhaps related to above, with prefix /č-/?

SIX: O: <hájo háikia>. Contains words for FOUR (in part) and TWO.

TO SLEEP. 1: B: <neianana> [nejanána] 'dormir'

TO SLEEP. 2: O: <îM> /íiM/. The only occurrence of <î> in the records.

SMOKE: O: <áanawa, ánawan>. Speaker unsure of this form.

SNAKE, SERPENT: O: <aúd>

SOLE OF FOOT see FOOT. 3 SOON see AFTER A WHILE

SOW see PIG. 3

SPANIARD see EARTH

TO SPEAK. 1: C: <aal> [a'ál] 'hablar'

TO SPEAK. 2: S: <gaxiamétët> in <upáat

gaxiamétët> 'long ago I used to speak'. This is unlikely to be Karankawa, but neither is it Tonkawa.

TO SPEAK. 3: S: <napé-nai pátsim> 'I speak, tell, converse'.

STAG: T: <tecomandotsen> [tëkomãdotsẽ] 'cerf ou chevreuil'. See DEER and BISON. 3

TO STAND: O: <jétso>. John McLaughlin compares Comanche /jïcï/ 'to fly', Panamint /jïcï/ 'to launch'.

STAR: C: <caguan> [ká(g)wan] 'estrella'

STOMACH. 1: B: <alouo, alouc> [a-luo, a-luk] 'ventre, estomac' = BELLY

STOMACH. 2: C: <enausa> [enáusa] 'estomago'.

TO STRIKE: O: <gá'an>

STRONG see MUCH

TO SUCK: O: <énno>

SUGAR, SWEET: O: <téskaus>. See MOLASSES.

SUN. 1 R /klos - klon/: T: <colone, colonee [sic]> [kolon, kolone] 'soleil'; B: <clos> [klos] 'soleil'; C: <clon> [klon] 'sol'.

SUN. 2 ? R /...owVl/: O: <dó'owal>. Possibly second part is the root found in MOON in Béranger, rather than the word for 'much' or 'strong'. Many Native North American languages, for example Algonquian languages, use the same word for 'sun' and 'moon'.

TO SWIM: O: <nótawa>

TAR: B: <couja> [kuža] 'goudron'

TATTOOINGS. 1: B: <bachénana> [bašenana] 'tatouages par piqûre'; cf. NEEDLE.

TATTOOINGS. 2: Gursky 1963: 32 lists <tampacuas> [tampákwas] 'lugar de los pintos o tatuados' (place of the painted or tattooed people) and <pacaguá> [paka(g)wá] as 'tatuado'. Tattooing was widespread among Texan tribes, including the Karankawa; it is believed that the Talon brothers were tattooed during their captivity, and Gatschet mentions that Sallie Washington bore faint lines on her face which were relics of tattooing during her time with the Karankawas, although the designs had long since worn away. The tribes of Southern Texas were often referred to, e.g. by Swanton, as Pakawan or Paikawan.

TO TEAR see TO BREAK

TEN: O: <dóotn hábe>. The first element seems to be DEER; the second is unidentified and otherwise unattested. It should be noted that animal names occur in the numeral systems of nearby languages, such as Atakapa ('hog') and Chitimacha ('rabbit'), where the terms are used in expressing 'hundred'. The Chitimacha term for 'rabbit, hundred' (/puup/) has in turn been borrowed as a loanword meaning 'hundred' into Natchez, where it is of course semantically opaque.

TESTICLES: C: <en> [en] '(t)esticulos'

THERE: O: <njá, niá>, also HERE.

THIGH R /...dal - del/: B: <em-edal> [em-ëdal] 'cuisse'; cf. semantic extension in C: <del> [del] 'toda la pierna' (the whole leg).

THREE: O: <kaxáji>

TIN see PLATE

TIRED: O: <kwá'al - kwál>

TOBACCO. 1 R /káhe/: T: <Cahé> [kaHé] 'tabac' (Swanton 1940: 126 reads <caké>), B: <a-caham> [a-kaHã] 'tabac', S: <káhe>, also <ka> in <ka swénas> 'cigarette'; C: <caje> [káxe] 'tobaco' (sic), also in <caje-tible> [káxe tíble] 'cigar'.

TOBACCO. 2: O: <dé>

TOES see FOOT. 2

TONGUE R /len, leen/: B: <aleane> [aleán] 'langue'; C: <len> [len] 'lengua'.

TONKAWA: S: <čankája>. Loanword from Spanish Tancahue? The Tonkawas' self-designation is different

/tickanwaatic/ 'genuine people' and unrelated.

TOOTH. 1 R /ej/: C: <ʔeg, ey> [eg, ej] 'diente'; O: <é>; also in <é tesselénia> 'toothbrush' (see BRUSH)

TOOTH. 2: B: <dolonaquin> [dolonakẽ] 'dent'. First part cognate with second half of <tequedolan> COW HORN and maybe FEATHERS.2. Original meaning 'protuberance'?

TORTILLA see BREAD

TO TOUCH: O: <čáutawal>. Gatschet thinks this is a compound <čauta + wal> (second part meaning MUCH?).

TREE. 1: C: <etsquequi> [etskéki] 'arbol'

TREE. 2: O: <akwiní>

TROUSERS: C: <? ṯenna-cama. [ténnakáma] 'pantalon'. Second part is widely-attested word for clothing.

TURKEY: C: <sam> [sam] 'keisolote' [sic, for Mexican Spanish *guajolote*], <sam-nen> [sam-nen] indicates 'female turkey, turkeyhen'.

TURTLE. 1: C: <ĉhaube> [šáube] 'tortuga' (turtle, tortoise).

TURTLE. 2: O: <háitnlokn> 'large green turtle' (an important source of food). Possibly a misunderstanding on O's part for a phrase meaning 'catch turtles' - the second half resembles Tonkawa /'óoxoloko/ 'turtle'.

TWO: O: <háikia>. Also in SIX, SEVEN, EIGHT, NINE and possibly FOUR.

UGLY WOMAN: C: <calee-bat> [kalé'e-bat] 'muger fea'; first half means WOMAN

TO UNDERSTAND see TO KNOW. 2

UTERUS, WOMB: C: <cuape> [kwápe] 'el útero'

VEINS: C: <? acuynu> [akújnu] 'venas'.

VERMILION: B: <cadum> [kadœ̃] 'vermillon'. An important trade article in the region.

VIRGIN: C: <ţetscuem-calí> [tetskwém-kalí] 'virgen'. Apparently or religious significance (occurring as it does between words for 'God' and 'Church'). Hardly from Spanish; just possibly connected with A BRAVE, in the sense of 'woman of chiefly or heroic rank' as the Blessed Virgin Mary may have been perceived by Karankawa neophytes. The second part is the word for WOMAN.

TO WALK. 1: B: <shacq> or <sotacq> [? s-hak] or [soták] 'marcher'

TO WALK. 2: O: <jé> = TO GO.

WALL: B: <enguesoul> [engësul] 'mûr'. Thus misread by some, notably Swanton 1940: 127; the reading and meaning mât being more likely. Mûr is not 'wall', which is mur (< Latin mūrus), but 'mature, ripe' (< Latin matūrus).

TO WANT. 1: O: <tá>. Also CAN

TO WANT. 2 see TO LOVE. 2

WAR: C: <? maĉhé, ?? traĉhé> [mašé, trašé] 'guerra'.

WATER. 1 R /klej/; B: <clay> [klej] 'eau'; C: <clé> 'el agua'; O: <gllé'i [sic], glé'>, also meaning 'sea, ocean, open waters'; also <téskaus gllé'> 'molasses', see SUGAR.

WATER. 2: T: <comcom> [kõkõ], 'eau', also <comcomdem> [kõkõdẽ] 'la mer'. This latter is a loanword from Atakapa, cf At. /káukau/ 'water', also Akokisa (far Western Atakapa) <cocaue> [koko] 'eau' (recorded by Béranger) and Caddo /kúuku'/ 'water' (Chafe n.d.), a loanword into Caddo which has been in use since at least the late seventeenth century: Talon records <coco> as the Caddo word for 'water' (other Caddoan languages using a stem */kiits-/). See Troike 1964 for details.

WATERHEN: B: <ouapa> [wapa] 'poule d'eau' TO WEEP: O: <owíja>. Gatschet thinks this was probably onomatopoeic.

WHERE (INTERROGATIVE ADVERB): O: <mudá>. Postpositional, thus <kíss mudá> 'where is the dog?'

TO WHISTLE: O: <áksool>

WHITE: S: <péka>

WIDE see BIG

WIND. 1: O: <bă, bë, bá>. Perhaps, according to Gatschet, a verb 'it blows'. WIND. 2: B: <eta> [eta] 'le vent'

WOLF: C: <badolú, badolú-nen> [badolú, -nen] 'lobo' (the form <badolú-nen> meaning she-wolf is not given a Spanish gloss). Apparently T: <quez>. B: <queche> 'chien' (dog) was also used for 'loup' (wolf).

WOMAN. 1: T: <achade, ? achadée> [ašad, ašade] 'femme'

WOMAN. 2 ? R /kad/lV/; C: <calí, calee-> [kalí, kalé'e> 'muger' [sic], also in VIRGIN, GIRL, PRETTY WOMAN, UGLY WOMAM. Maybe O: <káada> 'girl' is in some way connected.

WOOD. 1: B: <quesoul> [kësul] 'bois' (substance), also in MAST, DEERSKIN, RIFLE)

WOOD. 2: T: <cohal> [koHal, ? kwal] 'bois' , See BOARD?.

TO WORK: O: <takína>. Apparently a loan from Nahuatl tequipanoa /tekipánoa/ 'to work'; many US Indian languages have loanwords for the concept 'work', usually from Spanish trabajar, occasionally from arar 'to plough', though the Nahuatl word occurs in Yaqui (Eloise Jelinek, personal communication, 1991).

YES: O: <ië'ë, ihie'A, hié'ë, hi'iä, híë'A>. Interjection. Note also the chant at the religious festival where yaupon tea was consumed, as described by Mrs Oliver in Gatschet (1891:16/ 82): <Há'i-jah, há'i-jah [lento]; hai [forte], hái-jah, hái-jah, hái-jah [alto e basso, crescendo e diminuendo], hái! [fortissimo]>. YESTERDAY: O: <tuwámka>

YOU, YOUR(S), THOU, THY, THINE. 1: O: <áawa. áwa>

YOU. 2 ? R /VmV/; O: <m> in: <m' čá áwa?> 'how do you do?' (literally: how do you see yourself?). The parallel forms /n-/ and /m-/ for first and second person singular pronouns are widely attested in Native American languages. It is possible that the <em- emi- en-> prefix attached to a number of nouns in Béranger's vocabulary is a second person singular possessive prefix.

YOUNGSTER see BOY. 2

## Karankawa Texts and Sentences

Below are the attested utterances - the phrases, sentences and texts - of Karankawa, mostly collected by A. S. Gatschet. Each line is followed by a gloss of the Karankawa words, indicating their relative place in the English glosses of the vocabulary. No attempt has been made to impose Standard Average European grammatical interpretations on these textlets. The English translation, as given by Gatschet, is then added. Page numbers are given, coded to both sets of page numbers in Gatschet 1891.

In the case of the few stray forms not given by Gatschet, the original translation of the Karankawa, in the matrix language, French or Spanish, is given, and this is followed by an English translation of the sentence in the matrix language.

### Texts

These are not original Karankawa compositions, but translations of English nursery rhymes.

1)

Gatschet 1891: 81/145: (the original orthography has been preserved; many words in this rhyme would normally have long vowels).

<Nátsa kwáan kódn hákus akwiní
Tál áksol, tál áksol, tál áksol, ná čá
Nátsa kwáan glóëssën gás, gá'i dëmo'u
'N' čápn ódn áwa, hamála kwáan kódn!'>

The original English version runs thus:

'Little cock-sparrow sat up in a tree,
he whistles, he whistles and thus whistles he;
a little boy came with his bow and his arrow,
and said; 'I will shoot you, poor little cock-sparrow!'

<Nátsa kwáan kódn hákus akwiní
[One small bird sit tree]

Tál áksol, tál áksol, tál áksol, ná čá
[He whistle he whistle he whistle I see]

Nátsa kwáan glóëssën gás, gá'i dëmo'u
[One small boy come bow arrow]

'N' čápn ódn áwa, hamála kwáan kódn!'>
[I intend shoot you pretty little bird]

2)

This fragmentary rhyme occurs on the same page as the above. It is a variant of 'Cry Baby Bunting', as Gatschet observed.

<Ăhămmiš glós'n, kóM owíya,
áwa béhema gás mëssús.>

According to Gatschet, the original runs: 'Rockaby baby bunting, your father's gone a-hunting;/ mother's gone to get the skin, to wrap the baby bunting in.'

<AhÁmmiš glós'n, kóM owíya,
[Hush boy, not weep]

áwa béhema gás mëssús.
[You father come soon]

Sentences and Phrases.

These are taken sequentially from Gatschet 1891, in their order of first occurrence.

From Alice Williams Oliver:

[65/ 129]

<ne báwus kawá'i, ná'i dó'atn ahúk>
[I give horse, I deer kill]
'Let me have the horse; I have killed a deer'

<Karánkawa kóM tá takína>
[Karankawa not want work]
'The Karankawas do not like to work'

[73/ 137]

<gaí a dëmóa ná'i>
[bow and arrow I]
'The bow and arrow are mine'

<gaí a dëmóa áwa>
[bow and arrow you]
'The bow and arrow are yours'

<kóM ahájika>
[not friend]
'enemy; not friend'

<ná'i jé dó'atn ahók mësús>
[I go deer kill soon]
'I am starting soon to kill deer'

<kóM aknámus>
[not eat]
'not eatable, or, do not eat'

[74/ 138]

<ná'i amóak akwiní>
[I fall tree]
'I fell from the tree'

<ná'i amél, ta kwiamója aknámus>
[I hungry want bread eat]
'I am hungry, I want to eat bread'

<káada amóak, káada owíja>
[girl fall girl weep]
' The girl fell and wept'

hálba mušawáta takína, ašáhak kwá'al>
[chief always work now tired]
'the chief has worked continuously; now he is tired
(and) wants to sleep'

<ná'i áwa báwuus>
[I you give]
'I give you'

<Captain Jim áwa kosáta>
[Captain Jim you make]
'Captain Jim made it for you'

<áwa kanínma>
[your mother]
'thy mother'

<wól bá>
[much wind]
'strong wind'

<gás bá'ak>
[come house]
'to return home'

<bákta budáma wál>
[day gone much]
'day long past'

<ni báwuus tesnakwája>
[I give milk]
'give me milk'

<ná'i áwa kwiamója bawúus>
[I you bread give]
'I give you bread'

<hálba budáma, gás mëssus bá'ak>
[chief gone, come soon house]
'the chief has gone, he will return soon'

<glósEn káada dán>
[boy girl push]
'the boy pushed the girl'

<gás, glósEn!>
[come boy]
'come, boy!'
[75/ 139]

<ná'i glósEn kwátso>
[I boy ill]
'my boy is sick'

<téskaus gllé'i>
[sweet water]
'molasses'
[= 'sweet water']

<ná'i kóta kuwaí háitn>
[I hurry horse catch]
'I ran to catch the horse'

<ná'i béhëma háitn>
[I father catch]
'(go and) catch up with my father!'

<káada hákës bá'ak>
[girl sit home]
'the girl sits in the house'

<tál akwiní hamála>
[that tree pretty]
'this tree is pretty'

<tá îM>
[want sleep]
'he wants to sleep'

<ná'i jé medáu ódn>
[I go duck shoot]
'I am going to shoot ducks'

<ná'i jé dóotn ahók>
[I go deer kill]
'I am going to kill deer'

<ná'i jé wól>
[I go much]
'I walked considerably'

<ná'i áwa ka>
[I you love]
'I love you'

<gás, káada!>
[come girl]
'come, girl!'

[76/ 140]

<ná'i dëmóa káhawan>
[I arrow make]
'I make arrows'

<kwánnakwan akwiní?>
[grow tree]
'do they grow on a tree?/ on trees?'

<áwa katá; kaupn>
[you laugh say]
'you laugh! tell (why)!'

<kaupn ná'i béhëma gás bá'ak>
[say I father come house]
'tell my father to return home'

<ná'i kanínma béhëma tá káupn>
[I mother father want talk]
'my mother wants to speak to the father'

<áwa kAnínma klabán?>
[you mother healthy]
'is your mother well?'

<ná't kwáči kosáta mëssús>
[I fire do soon]
'I shall soon build a fire'

<ná'i kotá bá'ak>
[I hurry house]
'I am hurrying home'

áwa kwá'al! hákës!>
[you tired sit]
'you are tired! sit down!'£

<gáta kwán>
[cat small]
'kitten'

<ná'i kúM kwás>
[I not know]
'I do not know'

<áwa ná'i kwáss?>
[you I know]
'do you know me?'

<áwa kwátsu>
[you ill]
'are you sick?'

[77/ 141]

<glósEn aknámus kwiamója>
[boy eat bread]
'the boy is eating bread'

<ná'i áwa matákia>
[I you hate]
'I hate you'

<kíss mudá? áwan mudá?>
[dog where boat where]
'where (is) the dog? where is the boat?'

<n' čé áwa>
[I see you]
'I see you'

<në bawúus kwáči>
[I give fire]
'give me fire!'

<ná'i béx̂ma>
[I father]
'my father'

<ná'i gai>
[I bow]
'my bow'

<ná'i glósEn>
[I boy]
'my boy'

<kíss nótawa>
[dog swim]
'the dog is swimming'

<kíss niá>
[dog there]
'the dog (is) there'

<wál niá>
[much there]
'far off'

<ná'i čá awán njá>
[I see boat there]
'I see a boat over there'

<ódn dëmóa>
[shoot arrow]
'to shoot arrows'

<áwa ódn mEsúus>
[you shoot soon]
'shoot now!'
[= 'you may shoot presently]

<madóna aknámus plá>
[pig eat good]
'a pig is good to eat'

[78/ 142]

<gás! ná'i áwa tá!>
[come I you want]
'come! I want you'

<koM tá takína>
[not want work]
'he does not want to work'

<glósEn ém tá wól>
[boy want jump much]
'the boy wants to jump to a distance' (or 'far out')

<glósEn tá téskaus gllé'i>
[boy want sweet water]
'the boy wants molasses'

<ná'i tá hákës>
[I want sit]
'I want to sit down'

<ná'i ténno Walúpe>
[I also Guadalupe]
'myself and Guadalupe'

<glósEn aknámus ténno>
[boy eat also]
'the boy eats (of it) also'

<ná'i aknámus kwiamója téskaus gllé'i>
[I eat bread sweet water]
'I am eating bread with molasses'

<Col. Robinson téts'oa ahúk>
[Colonel Robinson cow kill]
'Col. Robinson has killed a cow'

<né báwuus kwáči! tólus! tólus!>
[I give fire run run]
'give me fire! run! run!'

<n' čá áwa>
[I see you]
'I see you'

<ná'i áwan čá>
[I boat see]
'I see/ perceive a boat'

<m' čá áwa?>
[you see you]
'How do you do?' (literally 'how do you find yourself?')

<n' čápn áwa ódn>
[I intend you shoot]
'I will shoot you'

<aM čutá>
[fish dangerous]
'octopus' (= dangerous fish)

<kóM aknámus tál aM; čúta>
[not eat that fish dangerous]
'this fish (is) not eaten; (it is) bad'

<wál glléi>
[much water]
'much water'

<éM wól>
[jump much]
'to jump to a (great) distance, to take a long leap'

<wál niá>
[much here]
'far off, way yonder'

<ná'i jé wol>
[I go much]
'I walked a good deal'

<ná'i béhëma wúak, tá îM>
[I father lie down want sleep]
'my father lay down to sleep'

Material from Old Simon (Gatschet 1891: 79/143)

<gaxiamëtét upáat>
[? speak long ago]
'long ago I used to speak [Karankawa??]'

<kwá má>
[horse black]
'black horse'

<kwán péka>
[horse white]
'white horse'

<xankí, níktam!>
[hurry young]
'come quick, boy!'

<napé-nai pátsim>
[I speak]
'I speak, tell'

<napé-nai naxerúaxa pára>
[I ? angry ? very]
'I am very angry'

<úši níktam>
[? man young]
'youngster, little man'

[80/ 144]

From Sallie Washington:

<ká'as wána!>
[come let-us-go]
'come here!'

From Rafael Chowell

<haĉhi cooĉhe>
'vamos a joder'
'let's copulate'

<mi-qüaĉhals>
'no te quiero'
'I don't like you'
(<mi-> + 'te'? <qüaĉhel> + 'querer').

## A Note on Karankawa Personal Names

By the time that Mrs Oliver observed the remnants of the Karankawas, in the 1840s, they had adopted Spanish or English personal names, for instance José Maria, Antonio, Walúpe (= Guadalupe), Lettie, or Captain Jim, and it is understood that they changed these names frequently to suit their own tastes; it is possible that a mortuary taboo existed concerning the names of the dead, as it did with the Tonkawas (Goddard 1979: 363), but we do not know for certain. The only Karankawa-language personal name which Mrs Oliver learend of was Kwáči, a male name which may mean 'fire'.

However, some earlier Karankawa names have come down to us from ther mission records of Nuestra Señora del Refugio and San Antonio de Valera, which were operative in the late eighteenth century, and these are cited with glosses in Johnson and Campbell (1992: 201); I mention them here for the sake of completeness of the Karankawa corpus. Thus we have the female name Equivec with a meaning 'arranged hair', and a male name Delençavan, which may be related to the root for 'tooth' (?? bright teeth, filed teeth). Some names belonging to the Coco, who spoke Karankawa, are also listed; these are the female name Cacatemelo 'dark woman', Eganandan 'short haired, bald', and Elencon 'eater of earth', the two latter names being male. These names serve to confirm the identity of certain Karankawa morphemes.

## Guy M. Bryan's Karankawa Vocabulary.

Early in the course of my correspondence with David J. Costa of the University of California at Berkeley, I mentioned my work on Karankawa and other languages of the Gulf Coast, and received from him some photocopied pages from the manuscript catalogue of the National Anthropological Archives, Smithsonian Institution, Washington DC, mostly relating to Karankawa. In this listing I discovered a source which had heretofore gone unmentioned in accounts of Karankawa linguistic materials, and I wrote to the NAA to obtain a copy.

The work listed was NAA MS 4944, a Karankawa vocabulary collected by Dr. J. O. Dyer of Galveston from Guy M. Bryan, 'an old patient', presumably during the first few decades of this century; it was suspected that most of the words were Spanish. I have no supplementary information on Mr Bryan or on where he might have learned his Karankawa.

On 23 August 1991, too late for the inclusion of the material in the body of the Karankawa vocabulary in my thesis, I received copies of Dr Dyer's notes with Swanton's annotations. I reproduce the substance of the material in this appendix, as its authentricity as genuine Karankawa is dubious. The copies themselves consist of two sheets of paper: on the first is a reproduction of the typed card of Karankawa words, with xeroxes of two pages appended; these two pages list a number of Coahuilteco words glossed into English. The other sheet has xerox photocopies of two other sheets of paper, one with Tonkawa words on it and the other has written on it a dozen random Karankawa words supplied with English glosses; these words are mainly taken from Gatschet 1891, with a few from Béranger or Talon: there is no new material here, and the words selected do not correspond to those on the typed card. The text of the typed card reads as follows:

'Words recorded by Dr. J. O. Dyer of Galveston, Texas,/ from Guy M. Bryan, 'an old patient', and perhaps Karankawa:/ He thinks most of them really Spanish:

aha, water
mego, friend (amigo)
cawa, horse (caballo)
mal, death
pawdo, lost
tato, farewell
muha, great

Caddo words noted by Dr. Dyer:
assi, leaves
bak, meat
maio, a turtle
copo, round
tongwa, a camp

Dr. Dyer says that the Karankawa called themselves/ Attapak' an'.'

As usual in Karankawa studies, these materials raise a number of questions to which answers are hard to find. The orthography of these forms is presumably a rough-and-ready one based on English spelling.

In regard to the Karankawa words, the words for 'friend' and 'horse' are Spanish loans attested in Mrs Oliver's materials, and 'horse' occurs in other vocabularies also.

The word for 'death' is recorded as <mal> 'dead' in Mrs Oliver's materials, while <tato> may be connected with her <ačáta>. I see <pawdo>
[? póodo] as a loan from some form of Spanish perder 'to lose'.

The word for 'water', <aha>, has Spanish agua added in ink, presumably by Swanton, but it may be a form of the Wanderwort for 'water' represented by more recent Tonkawa /'aax/ and found in Cotoname, later Comecrudo, Garza and Mamulique, and possibly reinforced by resemblance to Nahuatl /a-tl/ (from a Proto-Uto-Aztecan stem */pa-/). The word certainly does not resemble <comcom> or the word reconstituted as R /kley/.

As to <muha>, 'great', which does not resemble the other forms for 'great', /já'an/ or <counin>, I have no solution to offer.

If some elements in this list are also attested in Karankawa, I am not convinced that the list is wholly Karankawa - indeed only one term, 'death', is unambiguously Karankawa, with another possible one and three Spanish loans - and list the terms for the sake of completeness and because it has been labelled 'Karankawa'. Apart from the exceptions mentioned above, none of the surrounding languages which I have examined - Atakapa, Tonkawa, Caddo, Choctaw, Cotoname, Comecrudo - shed much light on the list.

As to the 'Caddo' words, the best that can be said is that they do not resemble the Caddo equivalents for the terms which I have been able to find. Some of them do, however, resemble Tonkawa words. The word for 'round' resembles the Tonkawa word for 'round' which Gatschet (1891: 82/ 146) gives as /kópol/ (Hoijer 1949: 33 /kopul/) and offers as an etymon for the Texan placename Kopano; there is a resemblance between Bryan's <maio> and Tonkawa /may'an/, 'land turtle' (Hoijer 1949: 13).. No parallel meaning 'meat' is available for <bak> (hardly Tonkawa /-ma'ek/ 'large cat'), while for <assi>, the nearest meaning to 'leaves' seems to be /'asoy-eylapan/ 'elmtree' (ibid.: 5).

The sense of <tongwa> as 'camp' may be an inference from the knowledge that the name for the Tonkawas (whose own name, /tickan-waatič/, meant 'genuine people [tickan]'), indeed the name which they used to refer to themselves when with other peoples, and which has passed to English and Spanish, supposedly derives from an unattested Waco (Wichita Caddoan) form tonkawéya, meaning 'they all camp together' (LeRoy Johnson, personal communication, November 1992).

As 'possibly Karankawa' terms one might append the placenames listed in Gatschet (1891: 81-82/ 145-146): Kopano, Aransas, Anaqua, Ecleto. They may be Karankawa, and they occur in former Karankawa territory, but they do not have any clear etymologies according to the Karankawa material which we have. But there again, as Gatschet points out, we also find Nahuatl, Tonkawa and Spanish placenames in the area. It is tempting to classify those naames above on historical evidence as Karankawa, when we have not a shred of linguistic evidence to support the claim; but this is not real classification, just sweeping difficult data under the carpet instead of admitting that in this case, as is so often true of the linguistics of Southern Texas, we shall probably never know its origin for sure.

## On the Use of Karankawa Data in Joseph H. Greenberg's 'Language in the Americas' (1987).

In June 1991 I was able briefly to examine Joseph H. Greenberg's controversial book Language in the Americas (Stanford UP, 1987). Having read a good proportion of the articles devoted to evaluating and criticising the work, I lost no opportunity in examining the Karankawa material incorporated in the book as part of Greenberg's contention that, aside from Na-Dene and Eskimo-Aleut, all the native languages of the Americas are ultimately genetically related in a huge family of several millennias' timedepth, called Amerind. I present my findings below.

## Analysis

I list below the Karankawa forms which he quotes, preserving his capital letters for glosses, and my evaluation of this analysis. They occur in two separate sections of the book: firstly, his treatment of 'Hokan', and secondly, his 'Amerind Etymological Dictionary'. Greenberg classifies Karankawa as being a separate branch in the Coahuiltecan subdivision of Hokan in the Amerind family, or to abbreviate this: Amerind C2c. I include page numbers, and the number and title of the 'cognate set' where the form is to be found.

The first set of forms comes from the section on Hokan (pp. 131-142).

p. 133: 8 ASHES Karankawa tap 'earth'. Not Karankawa: Coahuilteco tāp (Swanton 1940: 43)

p. 134: 15 BLACK Karankawa ma. Correct; recorded from Old Simon.

p. 135: 41 DOG Karankawa keš. Correct, and evidently taken from Béranger (who has <queche>); but Coahuilteco keš, listed before the Karankawa form, should be kačōwa. In any case. keš may be a Wanderwort.

46 EAT Karankawa ama 'fish'. Wrongly spelt; Karankawa form is am(Chowell) or aM (Mrs Oliver).

p. 139: 107 NOSE Karankawa emai 'nose'. Wrong: Béranger's form is emay alouacq, of which the second part is comparable to Chowell's lóo, while the first part is some prefix. Greenberg has segmented the form, but taken the wrong segment as the word for 'nose'. He appends this form to a set containing such words as Karok jufiv and Washo šujeb.

The second and larger set comes from Chapter 4 (181-270) entitled 'Amerind Etymological Dictionary'.

p. 194: 37 BLACK Karankawa pal. Correctly cited from Mrs Oliver (who tags it with a question mark), though it may be a loanword from Comecrudo.

p. 199: 50 BREAST Karankawa kanin. Correctly cited, and taken from Mrs Oliver; this is evidently a loanword from Cotoname, and cannot be used in historical-comparative work oriented to proving genetic relationships.

p. 210: 79 DIE (!) Karankawa mal 'dead'. Correct; taken from Mrs Oliver.

p. 222: 118 FLY (!), noun Karankawa kamex. Wrongly interpreted; Béranger's camaje suggests a form such as [kamež] or [kamej].

p. 226: 128 GO (3) Karankawa je. Correctly cited, if <j> is [j]; taken from Mrs Oliver.

p. 226: 137 HAND (1) Karankawa makuel 'five'. No such form exists in Karankawa (where 'five' has been recorded as /náatsa béhema/; makuél is 'five' in Comecrudo (Swanton 1940: 80), where it is a loan from Nahuatl macuillil.

p. 243: 178 MAN (2) Karankawa ahaks. This is an interpretation of one possible reading of a word in Béranger's vocabulary, written <ahax>, which might also be read as <alax>.

p. 260: 248 SUN (1) Karankawa auil 'moon'. An interpretation of a form from Béranger, which may be either <a-ouil> or <a-ovil>.

p. 266: 267 WHITE (2) Karankawa peka. Correctly spelt; a form taken from Old Simon.

p. 269 WING Karankawa hamdolak 'feather'. Wrongly spelt; Béranger's form is hamdolocq.

At least one Karankawa form is mislabelled:

p. 142: 153 TWO 'Tonkawa' haikia is Karankawa; Tonkawa has ketay for 'two'.

## Conclusions

It is apparent that Swanton 1940 has been Greenberg's sole source for Karankawa, and that he has not used the vocabulary in Landar 1968, much less the microfilm of the Berlandier and Chowell vocabularies.

Of Greenberg's fifteen 'Karankawa' forms, ASHES and HAND are not Karankawa at all, while the wrong portion of a form has been used in NOSE. Three forms - FISH, FLY, WING - have been wrongly spelt. The forms of MAN and MOON are open to other interpretations. The two words for BLACK, BREAST, DIE, DOG, GO and WHITE are spelt an analysed correctly and unambiguously. However, it is possible that one word for BLACK, BREAST, and DOG, are loans into Karankawa. In addition, a Karankawa word is miscited as a Tonkawa word.

In short, of sixteen real or supposed Karankawa entries, six contain mistakes of assignment to wrong language, misspelling or inaccurate segmentation, or forms reinterpreted without explanation, practices which characterise Greenberg's book as a whole (see for example Campbell 1989). The quality of Greenberg's work on Karankawa is appreciably higher than that for other languages, but it is still poor.

## REFERENCES

IJAL stands for International Journal of American Linguistics.

Berlandier, Jean-Louis. 1828-1829 [1969] The Indians of Texas in 1830. Edited and translated by John C. Ewers. Washington: Smithsonian Institution.

1980. Journey to Mexico during the Years 1826 to 1834. Translated by Sheila M. Ohlendorf. 2 vols. Austin: Texas State Historical Association and Center for Studies in Texas History, University of Texas.

Berlandier, Jean-Louis, and Rafael Chowell. 1828-1829. Vocabulaires de diverses peuplades nomades des parties septentrionales du Mexique, Lipans, Comanches, Tancahueses, Carancahueses, Carissos.. British Museum Additional Manuscripts 38720. [Formerly Phillips MS 15588. A set of unpublished vocabularies from South Texas and Northeastern Mexico, containing materials on Tonkawa, Karankawa, Lipan Apache, Comanche - the earliest attestations of these two languages - Cotoname, Comecrudo, Garza, Mamulique, and Yucatec Maya.]

Chafe, Wallace L. n.d. Caddo Word List. Unpublished manuscript, 8 pp.

Crawford, James M. 1988. On the relationship of Timucua and Muskogean. In: William Shipley (editor), In Honor of Mary Haas, pp. 157-164. Berlin: Mouton de Gruyter.

Dorsey, James Owen, and John Reed Swanton. 1912. A Dictionary of the Biloxi and Ofo Languages. Bureau of American Ethnology Bulletin 47. Washington: Government Printing Office.

Dyer, Dr. J. O. 1923. The Pioneer Texas Farmer, Galveston Daily News, June 3, 1923. (One of a series of articles on early Texan settlers and their encounters with Indians, written by a Galveston physician and based on interviews with aged settlers and other patients.)

Gatschet, Albert Samuel. 1886. Die Karankawa-Indiane. Globus (Braunschweig) 46 (8): 123-125.

1891. The Karankawa Indians, the Coast People of Texas. Archaeological and Ethnological Papers of the Peabody Museum, Harvard University 1 (2): 65-167 (also issued as a separate, i-viii, 9-103; both paginations are given in the text. The separate was subsequently reprinted: Gatschet n.d., n.p.).

Goddard, Ives.1979. The Languages of South Texas and the Lower Rio Grande. In: Lyle Campbell and Marianne Mithun (eds.), The Languages of Native America, pp. 355-389. Austin: University of Texas Press.

Grant, Anthony Paul. 1990. The Earliest Caddoan Vocabulary. MS, 8 pp.

Gursky, Karl-Heinz. 1963. Die Stellung der Sprachen von Nordost-Mexiko und Süd-Texas. Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft, Band 4. Nortorf (Germany).

Hoijer, Harry. 1949. An Analytical Dictionary of the Tonkawa Language. University of California Publications in Linguistics 5 (1): 1-74.

Johnson, LeRoy, and Thomas Nolan Campbell. 1992. Sanan: Traces of a Previously Unknown Aboriginal Language in Colonial Coahuila and Texas. Plains Anthropologist 37 (140): 185-212.

Landar, Herbert. 1968. The Karankawa Invasion of Texas. IJAL 34: 242-258.

Newcomb, William W., Jr. 1961. The Indians of Texas. Austin: University of Texas Press.

Pitkin, Harvey. 1985. A Wintu Etymology: 'Love'. IJAL 51: 536-538.

Powell, John Wesley. 1880. An Introduction to the Study of Indian Languages, together with a list of Words, Phrases and Sentences to be Obtained. Washington: Government Printing Office.

Robinson, Lila Wistrand, and James Armagost. 1990. Comanche Dictionary and Grammar. Norman, Oklahoma: Summer Institute of Linguistics.

Salinas, Martín. 1990. Indians of the Rio Grande Delta: Their Role in the History of Southern Texas and Northeastern Mexico. Austin: University of Texas Press. (Texas Archaeology and Ethnohistory Series).

Schaedel, Richard P. 1949. The Karankawa Indians of the Texas Gulf Coast. Southwestern Journal of Anthropology 5: 117-137.

Shaul, David Leedom, Katherine Turner, and James Collins. 1984. Esselen Linguistic Materials. Kansas Working Papers in Linguistics 9: 127-140.

Swanton, John Reed. 1940 Linguistic Material from the Tribes of Southern Texas and Northeastern Mexico. Bureau of American Ethnology Bulletin 127. Washington: Government Printing Office.

Taylor, Allan R. 1963. The Classification of the Caddoan Languages. Proceedings of the American Philosophical Society 107: 51-59. 1976. Words for 'buffalo'. IJAL 42: 165-166.

Troike, Rudolph C[harles}. 1964. The Caddo Word for 'Water'. IJAL 30: 96-98.

1985. Exorcising a Karankawa Ghost. IJAL 51: 599-602.

1987. Karankawa Linguistic Data. In: Robert S. Weddle (editor), Three Primary Documents: La Salle, the Mississippi, and the Gulf, pp. 288-301. College Station (Texas) Texas A&M University Press.

Villiers du Terrage, Marc de, and Paul Rivet. 1919. Les indiens du Texas et les expéditions françaises de 1720 et 1721 à la 'Baie Saint-Bernard'. Journal de la Société des

Américanistes de Paris 11 (2): 403-442.

1929. Deux vocabulaires inédits recueillis au Texas vers 1688. journal de la Société des Américanistes de Paris 21 (2): 307-311.

Wolff, Thomas. 1969. The Karankawa Indians: Their Conflict with the White Man in Texas. Ethnohistory 16: 1-32.

# RD-LEVEL NOMINALIZATION IN CHOCTAW

Marcia Haag
State University of New York at Stony Brook
University of Oklahoma

Abstract: Choctaw has a spare derivational morphology. This work discusses the limits of conversion, or zero-derivation, as a nominalization process, and shows that neither conversion nor a characteristic NP stress pattern are word formation rules. The prefix naa is a nominalizer that conforms to the theoretical constraints that both the base and the product of a word formation rule be predictable.

## Introduction

Little study has been directed toward nominals in Choctaw. Because the Choctaw verbal system is quite complicated, most morphological studies have tended to concern themselves with the verbs and their inflectional affixes and clitics.

In this work I will begin to fill the descriptive void and to situate the resulting description within current morphological theory. A framework of theoretical assumptions will be followed by a review the relevant literature to point out the areas of conflict in analyzing the nominalization process. I will concentrate on the derivation of nominal forms from other categories and will treat, of necessity, other seemingly unrelated problems, to wit: the status of a phonemic glottal stop, the status of adjective as a lexical category, pitch accent, and word boundaries.

Besides linguistic literature, great reliance was placed on the speech and judgments of native speakers. The reader should keep in mind that native speakers available to me had not been formally schooled in Choctaw grammar and spoke a variety of dialects whose relationships have not yet been described.

The goal of the study is to find true morphological processes in creating nominals. I will insist that candidates for nominalizing processes should be able to qualify according to theoretical criteria outlined later. Periphrastic solutions and syntactic solutions will be exposed as such but not analyzed further. Lexical processes such as compounding will not be examined.

*Kansas Working Papers in Linguistics,* Volume 19, Number 2, 1994, pp. 57-85

The primary results are that, first, Choctaw relies on lexical conversion, often termed zero derivation, for many nominalizations, but this is a process occurring in the lexicon and therefore not a word formation rule, and secondly, there is one reliable lexeme-level nominalizing affix, naa- .

Theoretical Framework. In limiting the problem of nominalization, I rely on some well-accepted definitions in morphology, particularly those of Aronoff (1976), Szymanek (1989), and Anderson (1992). I also draw upon Drijkoningen's (1989) work on the syntax of affixes to establish the syntactic levels of some particularly opaque constructions. I will leave out lexically-based word-formation strategies such as compounding. What we are after here is morphological derivation, with a word-level base and a word-level product at the end of each derivational cycle, if there are more than one.

The chief object of the search, then, would be a Choctaw word of unambiguous lexical category, that undergoes some overt phonological operation to produce a new word of predictable lexical category (in this case Noun) and whose meaning can be reliably if not exhaustively determined by the meaning of the base as a function of the word formation operation. In the happiest case, this function, or word formation rule, would also be productive, counting a large number of existing Choctaw words and permitting formation of new words (meaning only `unknown to the speaker') and even nonce words.

The simplest evidence would consist of a series of affixes (by definition not able to exist independently) which could be correlated with a lexical category, and with certain constraints upon the base form, concatenated in a way we are very familiar with in English:

$\text{convolute}_V$
$\text{convolution}_N$
$\text{convolutional}_A$

By way of fair warning, this will not prove to be the case in Choctaw. Lacking sets of recognizable and concatenable affixes, other means must be found that may produce the same effect. The danger here lies in seizing upon any process that appears to produce a syntactic noun without regard for type and predictability of the effect, and without testing for the employment of the same process in accomplishing other possibly unrelated tasks.

The Literature. Nominalization as a morphological process has received scanty attention in the general body of linguistic research on Choctaw. Previous treatments group around two general hypotheses. The first, incompletely described by Byington and again by Downing, suggests that there may be a rule permitting general lexical conversion. The second, attributable to Nicklas, involves (primarily) placement of an accent (undescribed) on the vowel of the penultimate syllable of the verb undergoing nominalization. A variation by Ulrich posits glottal stop in conjunction with penultimate accent as a nominalizer of verbs.[1]

In Downing's pedagogical sketch of Choctaw (1974), he suggests (p. 23) that Choctaw words regularly change lexical category without morphological encumbrance, that is, by zero derivation. He quotes Byington's manuscript notes to his editor:

> ‘Cannot all Choctaw nouns be treated as verbs? The root may be considered as in the infinitive mood: as *hattak* "to be a man"; *hattak* (with last syllable accented), "it is a man"; *hattak okmat* "if a man".'

(According to Downing, Byington's editors pointed out that real-world concrete objects are generally considered to be nouns primitively.)

Byington, in his Choctaw grammar (1870), makes a number of assertions as to the fluidity of lexical category membership, suggesting without arguing for a rule for general conversion:

> (p. 49) ‘The words used as adjectives, or attributes of nouns, are in reality verbs;' (p. 52) ‘Adverbs in Choctaw are verbs as well as adverbs;' (p.45) ‘The nouns are either primitive or derived...the latter are derived from verbs, adjectives, other nouns, etc.'

In the last case, ‘derived nouns', Byington lists a number of converted items, but also says (p. 47):

> ‘Abstract nouns are usually formed from neuter verbs as *kvllo* "to be strong". The translators of the New Testament rarely, however, use these words alone, but combine others with them, for example *nayimmi* "faith",...*nanisht i hullo* "love"...'

Nicklas (1972 p. 27; Jacob, Nicklas, and Spencer 1977 p. 144), states that some derivations of nouns from verbs and adjectives occur by the placement of an accent on the penult of the original word. He states that the meaning of the derivation (verb to noun in this case) produces an actor in the case of active verbs and a patient in the case of passive verbs (which are also derived).

Thus, from taloa `to sing' we get talóa `singer' and from holhponi `to get cooked' we get holhpóni `cooked food'.

The placement of penultimate accent seems to be a nominalizer of adjectives as well as verbs, according to Nicklas. Adjectives taking the accented penult name the abstract quality of their attributes: weki `heavy' becomes wéki `weight'. (In Nicklas's transcription, <e> = <ii> of my transcription.)

Furthermore, (same reference), nouns are also converted into other nouns by accent of the penultimate: chokfi `rabbit' becomes chókfi `sheep'. Nicklas groups these derivations together as examples of the same phenomenon.

Nicklas does not discuss what happens to verbs which already bear pitch accent (according to him), such as ikhána `to know'.[2]

From the above it would seem that accenting of the penultimate results in an array of effects, one of which may be agentive derivation (and its counterpart, the patient) in verbs, and another which may derive the name of an abstract quality in adjectives.

But it is clear that many other possible types of nominalization are left out, for instance, the simplest one, Nomina Actionis (to use Szymanek's term) wherein the verb is transformed in some regular way to mean `act of Ving'. Another type of interest to us here is the `single act of V'. The literature and my own investigations show there may be no morphological distinctions among these usages.

Nicklas's noun-to-noun derivations appear to be entirely lexical.

Looking now at Ulrich's work, his main claim is that `the only category-changing derivational morphology in Choctaw is the nominalization of verbs by accentuation of penult and suffixation of a glottal stop' (1986 p.77). He re-states this position in his 1993 work (p. 440): `The productive method of nominalizing verbs involves a glottal stop in the final syllable, together with an accent on the penultimate syllable in Choctaw...' To investigate, we must first wonder why there are two (for him) phonemic processes at work and if they must occur in tandem: this will mean discovering if either or both of the processes is indeed phonemic and differentiating them if such is the case. Ulrich does not discuss what kind of noun his process produces. The lack of any constraints on the output's noun type (gerund, agent, abstract, etc.) would seem to limit the usefulness of this process as a word-formation rule. Examples from his work show (1993 p. 443) chōpah `to buy' nominalizing to chōpa `buyer' or `purchase' (either agent or patient), and (p. 435) hilhah `to dance' nominalizing to hílha `a dance' (single act of V).

A second claim of Ulrich that bears upon the problem is that Choctaw has no lexical category Adjective (1986 p. 15). Since the Choctaw copula has a dubious status (Broadwell 1990; Davies 1986), and the usual construction is to inflect adjectives to produce stative predicates, Ulrich claims, echoing Byington's intuitions, that there is thus no separate category--`adjectives' are either inflected verbs or nominalized by accent and glottal stop. His lengthiest discussion is about attributive adjectives being in actuality `nominalized verbs' (1986 p. 78). He does not discuss the nominalization of other kinds of verbs (transitive and intransitive or unergative) as such, but does offer examples of nominalization of these types of verb occurring in the same way (discussed in Section 2).

Ulrich's argument depends upon there being a phonemic final glottal stop and accent on the penult of himitta that will distinguish (1) and (2). In Ulrich's gloss:

(1) sa- himítta?
lacc young nom
`I'm young' (Ulrich 1986 p. 79)

(2) sa- himitta-h
lacc young tns
`I'm young'

Ulrich's example in (1) seems to denote a syntactically free-standing utterance, and to make it more nominative-sounding to the English ear, might be better rendered `I'm a young one' or even `a young one, I'.

Broadwell (1990 p. 112) agrees that the lexical category Adjective (along with Quantifier) is a `subclass of verb', citing the display of verbal morphology on both lexical types. He provides a syntactic argument involving nominalization of this type verb through its appearance under a node that does not contain INFL, to give the briefest of summaries. I will offer another point of view in the next section.

As for nominalizing affixes, Byington's dictionary (1915) but not his grammar cites na and nan as nominalizers. The grammar also has a small list of words employing the suffix ka or kaka that seems to nominalize adjectives or what he calls `neuter nouns', that is, attributive adjectives that can be inflected as verbs; the resulting noun has the meaning `thing with attribute A'.

(3) chukbi `to be a corner' --> chukbika `a corner'

chito `big' --> chitokaka `God, the
(Byington, 1870; p. 45)

The suffix ka is rare in modern Choctaw.

Byington's grammar has this to say about (my spelling) naa (p. 47): `The suffix (sic) nana or nan gives an intensive signification--naniihollochi "accursed thing".' Clearly, this is not in keeping with his dictionary definition.

Nicklas (Jacob, Nicklas, and Spencer 1977), but not Ulrich, cites the use of naa- (p. 152):

> `Sometimes nan (na before consonants) is used to derive a noun naming the patient: nan ishko "a drink"...Sometimes it names the actor: na tolubli "jumper"..." In the next paragraph: `When the noun contains isht, it names an instrument: isht tiwa "key" (to get opened)...'

Nicklas does not treat the status of the two strings, naa and isht, but from his description naa clearly is derivational and isht does not appear to primarily affect lexical category.

To summarize previous treatments of the nominalization problem, we see that (1) Choctaw may have a general rule permitting verbs and adjectives to be nominalized simply by using them as nouns; (2) an accent placed on the penultimate syllable of a verb may be a nominalizer (per both Nicklas and Ulrich), (3) a glottal stop suffixed to the verb stem may be a nominalizer, either in concert with an accent as in (2) or alone (Ulrich), (4) there may be words, clitics, or affixes such as naa and isht that are nominalizers, and (5) there is no predictable semantic outcome for any nominalized verb or adjective claimed by any author.

### Preliminary Work: The Glottal Stop, The Adjective, The Accent, and Boundaries

The Glottal Stop. Ulrich claims a phonemic glottal stop in Choctaw in both his 1986 and 1993 work. The later work discusses Western Muskogean and contrasts Chickasaw/Choctaw glottal stop distribution, notably the presence of final glottal stop after a vowel in Choctaw where Chickasaw has only a vowel. Chickasaw, however, has medial glottal stops where Choctaw employs other phonological phenomena such as vowel lengthening (and see Footnote 3). Since Ulrich continues to maintain glottal stop suffixation as a productive nominalization rule in Choctaw, the glottal stop must have phonemic status.

In comparing the different distribution of the glottal stop in Chickasaw and Choctaw, Munro (1987 p. 120) finds ‘phonetic glottal stops only in final position,' in Choctaw. Broadwell follows Ulrich's view that (1990 p. 16): ‘Final glottal stop is added by rule to all final vowels.' He suggests that Ulrich may be correct in stating that glottal stop is phonemic, but for reasons that are convenient in solving the distribution problems of final [h] (pp. 18-20) rather than by independent evidence of phonemic status.

Looking to more traditional diagnostic tools for detecting phonemicity, there are no known minimal pairs involving glottal stop, and none of my consultants can recognize the glottal stop.

My consultants do employ what may be a word-final glottal stop allophonically; however, they perceive themselves as simply terminating a short vowel, especially when showing contrast with final /h/. I have noted a good deal of variance among speakers: some produce a clear truncation of final vowels that is undifferentiable from a phonetic glottal stop; others use it variably, and some not at all. It is not unreasonable to suppose that the ‘glottal stop' in the latter case is really the sensitivity of the English ear to the impossibility of a final non-low non-diphthong, per Sapir's treatment (1925).

(4) [Hattak-a? balili-h]
man subj run pred
‘The man is running.'

(5) [Ofi? pisa-li -h]
dog see 1nom pred
‘I see a dog'

I can find no support in my own field research for a phonemic glottal stop. I cannot directly answer some of Ulrich's arguments that depend on its presence, particularly his analysis of the glottal stop as a `post-lexical clitic' (1986 p. 107-109). However, since Ulrich seems to require that a penultimate accent accompany the glottal stop, I will hypothesize that this accent will do to nominalize, as Nicklas claims.[3]

Even if phonemicity cannot be established for glottal stop, a practical argument that may inadvertently promote this view regularly insinuates itself into the discussion, if not the literature per se, about Noun/Verb distinctions. That argument is based on the observation that if predicates are marked by /h/, and verbs are the quintessential predicate, lack of /h/ (and presence of final glottal stop whether phonetic or phonemic) marks a deverbal noun. Allen Wright's lexicon (1864), in contrast with Byington's dictionary, marks (most) predicates with /h/ and deverbal nouns without it. The danger here is to confuse a practical consequence with a word formation rule: as stated in the theoretical section, a productive word formation rule must constrain and predict both the base and the output. Deverbal nouns must perforce sound like nouns--generally /h/-less. The converse is not implied: /h/ deletion does not qualify as a word formation rule unless it can meet theoretical constraints, otherwise it cannot be distinguished from zero-derivation. As we will see later, predicative /h/ may be present or absent on members of other lexical categories.

The Status of the Adjective as a Lexical Category. The nature of lexical categories in Choctaw is a topic that is attracting in-depth research apart from this investigation. A considerable body of evidence exists for a category Adjective, some of which will be brought to bear on the argument at hand. Ulrich, and to a lesser extent Broadwell, asserts that Choctaw dispenses with adjectives as a syntactic category, utilizing only Noun and Verb as the major lexical categories. Ulrich's argument for verbal status is that adjectives can be inflected (1986 p. 15). Ulrich recognizes `nominalized stative verbs' (adjectives) by their bearing penultimate accent and final glottal stop.

Ulrich's examples include, in his gloss (1986 p. 78):[4]

(6) 0 Chaaha -h.
3acc tall v
`he is tall'
(7) Hattak chaaha' pisa-li-h.

man     tall   see   I v
'I saw the tall man'

Ulrich also suggests that verbs can be nominalized with their clitics:

(8) Ish-híła  -?  (p. 79)
you dance  n
'You're a dancer.'

Ulrich states that *chaaha?* in (7) is a 'nominalized verb'. Since both *hattak* and *chaaha* cannot both appear as nouns in the same Noun Phrase unless they are a compound (and he agrees they are not), he argues for the status of the glottal stop as a post-lexical clitic, nominalizing the clause *hattak chaaha*, which was deverbalized when it lost its inflective element *-h*. In the case of nominalizations incorporating clitics such as (8), only a clausal analysis is possible.

Recalling that the motivation for viewing adjectives as verbs was their capability for inflection, it is revealing to point out that quantifiers, degree specifiers, adverbs, and nouns may also be inflected, although nouns as predicate nominals may not generally bear the *-h* predication marker. All but nouns may also carry aspectual morphs, which are commonly seen on verbs (and which Ulrich assigns to verbal, but not nominal, morphology (1986 p. 15).) The presence of verbal morphology alone would seem to make for but a tentative diagnosis of lexical category, recalling that there is no evidence that Choctaw does not employ polyfunctional affixes.

The following examples show the wide distribution of the predication marker (*-h*), the complementizer *-kat*, and the stative aspect marker, all examples of typically verbal morphs.

(9) Soba balili-kat losa  -h.
horse run   comp black pred
'The horse that's running is black.'

(10) Soba toklo-kat balili-h.
horse two  comp run    pred
'Two of the horses are running.'

(11) Soba  losa- t   toklo -h.
horse black subj two  pred
'There are two black horses.'

(12) Soba hannali mat losa  -h   chiyyohmi-h.
horse six    dem black pred very     pred

`Those six horses are very black.'

(13) Soba-yat itōla -h.
horse subj lie/asp pred
`The horse is lying down.'

(14) Chokfi-yat chokka nōta itōla -h.
rabbit subj house below/asp lie/asp pred
`The rabbit is lying just below the house.'

A safer conclusion would be that Choctaw makes multiple use of a small number of inflectional affixes and verbal status cannot be determined from the appearance of an affix that coincidentally appears on verbs.

Turning to positive evidence for a category Adjective, following Chomsky's (1965; 1970) characterization of Adjective as [+N, +V], we see that Choctaw `adjectives' also behave nominally: only adjectives and nouns may employ the verbal proform -a in one kind of focus construction. When so used, they denote permanent characteristics i.e., (Adj.) person, or may have metaphorical force.

(15) Chito si- a -h.
big 1acc V pred
`I am a big person. ~ I am an important person.'

(16) Alikchi si -a -h.
doctor 1acc V pred
`I am a doctor.'

Furthermore, `adjectives' are readily nominalized by a general conversion rule, producing `the A one.'

(17) chito mã `that big one'

ãchito `my big one'

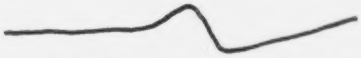
This latter rule will be investigated more throroughly in the next section, when I compare the behavior of this lexical group with that of (true) verbs.

There are more general grounds upon which to discourage the argument that an Adjective is variably a Verb or a Noun. Recall that this argument relies on the loss of tense (INFL) as a de-verbalizing process that then renders the former verb capable of being reconstituted as a noun. The main problem with this kind of subtractive derivation is that there must be sufficient information in the lexical residue to permit reorganization as a word that will function like all other adjectives and NOT like members of other `verbal subclasses'. Only lexical specification will do this; syntactic solutions depend on these differentiations primitively. We would have to propose some sort of lexical subgroupings to be sure that the verb decomposed into the correct category once inflection were lost--this is not different from having lexical categories to begin with.

A far more conservative solution is to approach the problem as one where most lexical categories in Choctaw are permitted to become predicates. Note that the output of such constructions is highly constrained: they render only predicates with the meaning `to be X', very much like English with its parallels `to be fat'; `to be many'. Furthermore, each lexical item behaves like its category mates with respect to both its behavior while serving as a predicate and its behavior freestanding. The traditional nomenclature and definitions obviate the need to posit clauses inside the Noun Phrase.

Accent. The problem of accent is central to this study. Both Nicklas and Ulrich suggest that the placement of penultimate accent on verbs and adjectives nominalizes them, at least some of the time. I would expect such a process to lend phonemic status to the accent (independent of any underlying lexically specified accent). To test this hypothesis, we need to differentiate among processes that would cause the voice to alter its pitch. Besides underlying `accent' assignment, (and leaving this notion not precisely defined for the moment) let's consider phrasal pitch contour (not treated anywhere in the literature).

Phrasal pitch contour: Choctaw speakers set off sentence constituents with vocal inflection. The phrasal contour for the simple declarative (three-constituent) sentence is Mid Tone on the first constituent followed by a High peak and low trough on the second constituent, and finishing with rising pitch the final constituent, usually the main verb:

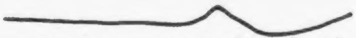


(Byington (1870) offers the following (p. 11):

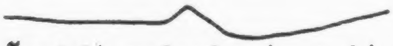
> `There is another accent which falls on the final syllable of such words as in English are followed by marks of punctuation, from the comma to the period. It is called the pause accent.'

This appears in keeping with the High finish, but since we do not know what Byington's `accent' is, this statement is at best something to muse over.)

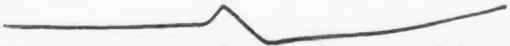
Note how this pattern is adhered to in each of these three-word sentences, even though the first two words of each represent different lexical categories or constituents, or even different clauses, as in (20).



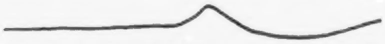
(18) Hattak-at tamaha ia-h.
subj loc V
`The man is going to town.'



(19) Ãnakfi palaska im-a-li-h.
IO DO V
`I give bread to my brother.'



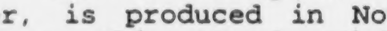
(20) Ohoyo-at sa-pisa-tokã, aachi-tok.
dependent clause V
`The woman said that (someone else) saw me.'



(21) Ohoyo-at taloa-h moma-h.
subj V Adv
`The woman is still singing.'

The contour seems to be positioned so that the main verb, or last word in the VP, as in (21), falls on the rising contour, no matter how many syllables it contains or if agreement markers are attached.

Choctaws are, of course, not obligated to this contour and may alter it to enhance expression. That this is the usual declarative contour may be demonstrated by its distinctiveness from the interrogative contour:



(22) Hattak-at tamaha iah-õ?
man subj town go interr.
`Is the man going to town?'

The falling pitch of the yes/no interrogative is characteristic.

Noun phrase contour: A somewhat weaker contour, in the sense of being more variable from speaker to speaker, is produced in Noun Phrases (Nicklas also attests to this (Jacob, Nicklas, and Spencer 1977 p. 13)). A noun phrase with modifiers will have each modifier accented on the penultimate syllable, unless it has underlying accent elsewhere. A noun phrase consisting of a lone noun may or may not bear penultimate accent.

(23) sóba `horse'

soba lósa `black horse'

soba lósa chíto `big black horse'

soba lósa chíto mã `that big black horse'

Notice that accent is shifted in accordance with NP boundaries, not word boundaries, even though *soba* `horse' contains an underlying accent according to Nicklas (Jacob, Nickas, and Spencer 1977). The same phenomenon occurs with Nouns that purportedly have no underlying accent.

While not offered as direct evidence for or against a phonemic accent, the phrasal pitch contour shows that a high (falling) pitch on a selected syllable of the last word of the constituent directly to the left of the verb complex is the expected pronunciation irrespective of lexical category, particularly if there is also a constituent to occupy the `initial Mid pitch position'. Additionally, noun phrases may bear high pitch on the penultimate syllable of the head if alone, and regularly on any adjectives and other non-determiner modifiers.

*Alternate lengthening:* In Choctaw, a series of CV syllables will undergo rhythmic lengthening (Ulrich (1986), or Alternate Lengthening (Nicklas 1974), whereby the second and every other even-numbered syllable except the final will be lengthened. Byington (1870; p. 11) offers an even more general scheme in which the penult and every other syllable moving leftward from it is also accented irrespective of weight. Lombardi and McCarthy (1991), drawing primarily on Nicklas's data, use this evidence to posit an iambic foot structure for Choctaw. Broadwell, too, notes that there is variation, whether dialectal or idiolectal, in the speakers' production of long vowels. His example, (1990; p.13) *sa-salaha* `I am slow' shows alternation between *sa-saalaha* and *sa-saalaaha*.) The relationship to accent is the same as that of contrastive vowel length.

*Contrastive vowel length*: Vowel length is contrastive in Choctaw. I am attempting to keep vowel length separate from the idea of `accent'; nevertheless vowel length and phonological stress are often strongly correlated in many languages. I bring it up here because Ulrich describes his `penultimate accent' occurring in expressions that (p. 78) `typically refer to characteristic activities or properties...íshko? "he drinks" or "he's a drinker".' The nomic tense (Nicklas's term) in Choctaw is made by lengthening the vowel of the penultimate syllable and denotes activities or states of being that are characteristic.

(24) Baryshnikov-at hiiła -na Sinatra-at talooah.
subj dance(asp) and subj sing(asp)
`Baryshnikov dances and Sinatra sings.'

It should be pointed out that empirical evidence for a nomic aspect is less conclusive than that for the other aspects. 5

Phonemic high falling pitch: Besides all the above, there is direct evidence for phonemic accent--this is rendered as high, falling pitch, contrasting with length, with euphonic word-final rising pitch, and with phrasal pitch contour.

The imperative is produced by placing an accent (high falling pitch) on the final vowel.

| | | |
|---|---|---|
| (25) | binili | `to sit' |
| | binilí | `sit down' |
| | tõshpa | `to make haste' |
| | tõshpá | `hurry up' |

The intensive aspectual marker is formed in part (along with medial segment reduplication) by high, falling pitch on a first-syllable vowel.

| | | |
|---|---|---|
| (26) | chito | `big' |
| | chíyyito | `huge' |
| | falama | `to return' |
| | fállaama | `to finally return' |

These distinctive processes will prove useful when testing penultimate accent as a nominalizer.

Word Boundaries. As previously mentioned, Choctaw has little in the way of a modern and well-disseminated written corpus. While this fact should not hinder the ability of an experienced linguist to pick out orally produced words from higher and lower levels of organization, it does tend to create disparities in what informants believe to be `words'. In the linguistic literature, authors tend to place hyphens between a root and its affixes and clitics. One pedagogical work (Jacobs, Nicklas, and Spencer (1977) separates nearly all morphemes. The main Choctaw texts, the Bible translated by Byington and a hymnal, attach some but not all inflectional particles directly to their hosts and frequently leave other elements that appear to be affixes or bound morphemes of some sort (notably *naa* and *ishtl*) free-standing. Compounds are sometimes written together and other times separated without appeal to clear phonological or semantic grounds. My consultants were not altogether helpful in this way either, since they were often sensitive to the meaning changes afforded by affixes and clitics and wanted them to be words (we might speculate about interference from English).

Since I am, from the above, clearly insisting that there should be affixes and clitics in Choctaw, it would be well for me to set out what sorts of criteria I expect to use to differentiate them from each other and from words.

Borrowing heavily from the work of others, especially that of Zwicky (1985), and Klavans (1985), I would use the parameters of independence (of an item) and attachment (preferences and requirements) to categorize candidates. The immediate importance to the problem at hand is that, of phonological entities realized as a string of one or more segments, a word-level nominalizer would best be an affix.

In accepting an entity as an affix, it should have a limited and predictable distribution (discussed earlier), should in some cases `close a word to further affixation,' (Zwicky's phrasing), should be affected by the syntax, and should NOT attach to an entire phrase. Our candidate *naa-* will be tested according to these diagnostics (not an exhaustive list).

## Three Hypotheses for Nominalization Processes

Lacking evidence for many overt derivational affixes, we should entertain the possibility that syntactic categories are derived by some conversion process. Certainly, Byington's Choctaw dictionary operates according to this theory, listing the same entry for related nouns, verbs, adjectives, participles, and adverbs:

| | | |
|---|---|---|
| tõshpa | `hasty' | A |
| | `speed' | N |
| | `to go in haste' | V |
| | `precipitated' | PP |
| | `speedily' | Adv |

Allen Wright, in his lexicon (1860), noted predicative uses of words, adding predicative *-h* to those forms that required it, such that he enters falamah `to return' but falama `a return'. (Of course, Byington, in describing usage, suffixed tense markers, although not -h, in at least those forms he used as predicates.)

Most words do not lend themselves to all categories, but it is extremely common for the same word to be used as both noun and verb, with the latter taking appropriate tense/predication marking.

Byington's dictionary does not take accent into account (although length is indicated), so accent may distinguish categories from one another.

It would seem we could hypothesize several solutions based on what facts we have and devise ways to choose the most convincing (if any).

Hypothesis 1. Nominalization occurs by conversion (zero derivation).

If we assume a basic verb (Szymanek p. 84) and propose that nominalization is verb-to-noun (and in Choctaw, adjectives are so readily made predicates, that we could posit an intermediary derivation of A --> V), then verbs are converted to syntactic nouns by a rule that operates in the lexicon. On the morphological level, such a noun cannot be distinguished from a verb (or adjective). What we don't know is whether Choctaw permits this rule to occur generally or on selected, and therefore, unpredictable items.

Such a hypothesis would be easy enough to test: we need only find verbs posturing as nouns. One problem we may encounter is differentiating nominal clauses from `words' and types of nominal clauses from each other; in these cases we will need to employ tactics that will illuminate syntactic level. Tests will include ability of derived forms to accept adjectives, determiners, case markers, and clitics.

A second crucial consideration is to predict a regular noun type output; while we might unreflectively expect this output to be a gerund or `condition of being V,' since the gerund represents a syntactic level different from the infinitive, this will have to be examined. Evidence from dictionaries seems to support the view that output type is variable and unpredictable: sometimes the noun type is an agent, sometimes a patient, other times `single act of V', etc. (see theoretical section).

Hypothesis 2. Verbs are nominalized by placing an accent on the penultimate syllable. Nominals do not bear predicative -h.

I am assuming that glottal stops are not nominalizers for the reasons discussed in Section 2, though a word may manifest its status as a noun by terminating in a phonetic glottal stop.

This hypothesis presents the most difficulty in testing. First, we must decide how we know if a syllable is accented, since CVC and CVV syllables are lengthened relative to CV, and some words have underlying stressed syllables, which surfaces as high falling pitch. Additionally, alternate lengthening operates at the word level, lengthening the vowels of some syllables. Then we must consider the effects of phrasal pitch contour. We might guess that the effects of a nominalizing accent might be obviated if the affected syllable is already `accented' somehow. Since we have a phonemic accent--high falling pitch--this is the only reasonable candidate for a nominalizer among this group.

To test Hypothesis 2, I select verbs with no attested underlying accent. I then ask subjects to pronounce sentences with these verbs in both a verbal and a nominal position. Then, on a different occasion, I ask subjects to listen while I pronounce the word IN ISOLATION both with and without penultimate accent. This last condition is critical in establishing a morphological, word-level effect. It is also important to look for the possible intrusion of conversion into this set of data; that is, if a speaker determines a verbal form to be a noun, we must assure ourselves that he is responding to the presence of the accent.

As for the type of the derived nominalization, since the authors supporting this rule predict a mixed bag of types, I will accept an array of noun types, although a robust rule should restrict the range of types.

Hypothesis 3. The prefix naa- is a nominalizer resulting in a derivation meaning `that which Vs/ is Ved.'

This hypothesis can be tested both in examining existing dictionary words and by creating hypothetical derivations by affixing naa- to a verb.

I test this by pronouncing both made-up and dictionary examples of this derived form in isolation to subjects and asking them what, if anything, they mean. Again, an effect at the word level is necessary to suggest a morphological nominalizer.

Responses are examined for consistency in the meaning of the derived words.

Other candidates for morphological nominalizing affixes must also be examined, particularly -isht- and aa- and its variant ai-

Results and Analysis. The results of tests support Hypothesis 1 as a lexical process but not as a word formation rule, and strongly support Hypothesis 3; Hypothesis 2 is not supported on several grounds, but bears discussion as to the relationship between accent and NP category as discussed in the previous section.

Conversion (Hypothesis 1): A very common means of derivation is lexical conversion. Some verbs have a nominal use that is given in the lexicon and whose meaning is unpredictable among of an array of noun types: \`single act of V'; \`actor'; \`abstract condition of V'.

In Choctaw we do not seem to have a nominalized verb comparable to the gerund. All clauses, whether tensed or not, contain predicative -h or a complementizer.

(27) Taloa-h sa- banna-h.
sing pred 1acc want pred
\`I want to sing.'

(28) Tobi impa-kat im- achokma-h.
beans eat comp 3dat good pred
\`He/she likes to eat beans.'

Drijkoningen (1989) devises a useful model for determining the syntactic level of nominalizations. The lowest level (and our research object) is the word-level (or $X^0$, or lexnom, in Drijkoningen's terminology). Intermediate levels would be gerund, at the verb phrase level, and infinitive, at the inflectional phrase level. Specifically, a gerund should permit the assignment of accusative case to an argument and permit the attachment of genitive markers, while an infinitive should not permit genitive markers.

Enlarging on Drijkoningen's model for our purposes, a lexical nominative ($X^0$) should behave as a lexical noun ($X^0$): It should accept adjectives, possessive markers, and casemarkers, and should also be able to accept tense and aspect markers (in contrast to the infinitive). Since infinitives are clearly marked as verbal, we need only concern ourselves with lexeme-level verbs.

Looking first at adjective acceptance, we see that some verbs but not others can be modified by adjectives.

(29) * omba achokma `good rain'

but hoponi achokma `good cook'

The subject marker used for lexical nouns, -at again, does not attach to converted nominals, except when these are lexicalized.

(30) * Taloat achokmah.
Sing subj good tns
`singing is good'

(31) lexicalized verb toksali `to work; a job'

Toksali-yat achokma-h.
work subj good pred
`The job is good.'

Possessive markers may similarly attach to converted nominals only when lexicalized.

(32) * Ā̃- taloa ish-hăklo-h-ō̃?
1poss sing you hear int
`Do you hear my singing?'

Chī-tõksali pi- banna-h.
2poss work 1acc/pl want pred
`We want your job.'

There is clearly no rule of general conversion from verb to noun in Choctaw. Still, a fairly large number of verbs are lexicalized, and as is common to such lexicalizations, without predictable outcomes as to type (the following are from Byington's dictionary):

| | |
|---|---|
| (33) hoponi `to cook' | hoponi `a cook' |
| komonta `to be uneasy' | komonta `uneasiness' |
| kocha `to come out' | kocha `a departure' |
| kocha `an outcast' | |

(and many others. The basic meaning is the postposition `out', the one my consultants select first.)

Again, the large and unpredictable array of derived noun types is a hallmark of lexicalization by conversion, and not evidence for a word formation rule.

<u>Converted adjectives:</u> A different story is told with adjectives undergoing conversion to syntactic nouns, first mentioned in Section 2. There is no problem with these accepting determiners and case markers, possessive markers, or degree specifiers and numerals. The output is always Adjective qualifying a null Noun head: `Adjective one'.

| | | |
|---|---|---|
| (34) | chito mã | `that big one' |
| | chito chiyyohmi | `the very big one' |
| | ã-chito | `my big one' |
| | chito-at | `the big one (subject)' |

This evidence suggests that there is a category Adjective that is differentiable from the category Verb.

Accent (Hypothesis 2). The presence of a NP pitch contour means that a noun could be differentiated from a verb on the basis of an accent on the penultimate syllable of the former. It is important to recognize that NP pitch contour is the result of another process--the conversion of a Verb Phrase to a Noun Phrase through zero derivation--and is not itself the word formation rule that derives a noun from a verb.

The following group of verbs does not have a corresponding set of nominals. Whether Choctaw speakers heard them pronounced with even stress or with a penultimate accent, they always stated that each word was a verb.

| | | | | |
|---|---|---|---|---|
| (35) | impa | `to eat' | ímpa | `to eat' |
| | haklo | `to hear' | háklo | `to hear' |
| | taloa | `to sing' | talóa | `to sing' |

As it fails to produce a word-level effect, we must withhold support for the penultimate accent as a morphological nominalizer. Since this experiment in eliciting judgments is not controlled for any kind of bias, we cannot be sure of what the consultants were responding to, but we must remain suspicious of the accent's lack of effect.

<u>Naa-</u> and Other Affixes. The prefix <u>naa-</u> regularly nominalizes verbs pronounced in isolation and the derived forms predictably mean `that which Vs/is Ved'.

The words in the previous example do not have a lexicalized *naa-* form. When they were derived as nonce forms, informants responded this way:

(36) nanimpa `food'

naahaklo `something that is heard'

naataloa `a song'

All these examples happen to have a patient role. It may be that this reading is the more productive, since the informants could have defined, but didn't, *naahaklo* as `a hearer' and *naataloa* as `a singer'. Byington's dictionary contains many agent readings, and in fact defines *naahaklo* as `hearer'. There does not seem to be at this time any reason to think that both are not licit, but I do not know what the distribution might be.

*Naa-* is also hugely productive in Choctaw, as in this monolingual speaker's invented word:

(37) naa-kapassa-chi
nom cold cause
`refrigerator'

If we look at another morpheme, *-isht-*, suggested by Nicklas as serving a parallel function to that of *naa-* we see that while *-isht-* adds the meaning `instrument' to the construction, it does not nominalize (or perform any other derivation), although its attachment may create a new lexicalization that is a de facto category change.

(38) ishko `to drink' ishtishko `a cup'

ia `to go' ishtia `to carry'

mĩti `to come' ishtmĩti `to bring'

Compare these with what happens when we add *naa-*

(39) apiisa `to measure'

ishtapiisa `to measure with'

nanishtapiisa `measuring instrument; ruler'

There are other morphemes that coincidentally form nouns in the course of a derivation. *Aa/ai*, a locative morpheme, frequently renders nouns: *aiimpa* `place where one eats' --> `table'. But *aa/ai* primarily confers the sense of location: *aiahoba* `to appear (somewhere)', and includes an enormous number of verbs that incorporate aa/ai with the sense `from', such as aatoba `to be made from'.

## The Nature of Naa-

That naa- nominalizes cannot be doubted, but there remain questions about its categorial status. It is not certain that it is an affix, though I treat it as one here and hope to offer persuasive evidence that it is an affix rather than a clitic or a word.

Ulrich states (1986 p. 16) that naa is a clitic, grouping it with isht and a number of other morphemes, none of them derivational in the sense of changing lexical category. As we recall, he confines nominal derivation to a accent-and-glottal-stop operation, and does not entertain affixal possibilities.

Byington defines naa (in his spelling na) as a prefix that attaches to verbs and nominalizes them, but he only occasionally attaches na orthographically to his entries; generally it stands alone. (The same treatment occurs with isht.)

Let us proceed with a number of familiar tests.

Susceptibility to Further Derivation. We would expect that an affix would lend a word limited permeability by other affixes, that is, in its derived form, a word then serves as a preferred base for another affix. Since there will be no other affixes, we can examine naa-derivations' behavior with respect to clitic attachment and conversion into verbs.

Naa derivations can be re-verbalized.

(40) fohka `to put on/in'

naafohka `clothes'

naafohka + tense `one is (tns) dressed'

naafohka fohka `to put on clothes'

(41) Ohoyo-at naafohka fohka -h.
woman subj clothes put on pred
`The woman is putting on clothes.'

(42) Ohoyo-at naafohka-h.
woman subj clothes pred
`The woman is dressing.'

Re-verbalized nominalizations can accept aspect markers.

(43) Ohoyo-at naafõhka -h.
woman subj be dressed/asp pred
`The woman is (stative) dressed.'

<u>Naa-</u> derivations can take adjectives, numerals, determiners, and casemarkers.

(44) Naafohka himona pã chompa-li -tok.
clothes new dem buy 1nom pst
`I bought these new clothes'

(45) Naafohka-at himona-h.
clothes subj new pred
`The clothes are new'

(46) Nanikhanachi toklo-kat sa -haklo-tok.
nom teach two part 1acc hear pst
`Two of the teachers listened to me.'

<u>Naa-</u> cannot attach to the clitics of inflected verbs, although there are a number of lexicalizations that include, especially, dative markers.

(47) *nan-im -achokma-h
nom 3 dat good pred

* naa-chi -pisa-tok
nom 2 acc see pst

but nanittimapiisa `contract'
from nan + itta + im + apiisa
nom + `together' + `3 dative' + `judge'

<u>Naa-</u> cannot attach to the possessive marker.

(48) * nan-ĩ -chokka
nom 3 pos house

<u>Naa-</u> cannot attach to phrases or clauses.

(49) * naa tobi impa-h
nom beans eat pred

From this evidence we can infer, using Zwicky's diagnostics, that <u>naa-</u> behaves like an affix because `syntactic rules can affect affixed words but cannot affect clitic groups', and `clitics can attach to material already containing clitics but affixes cannot,' (1985 p. 285).

Ordering and Exclusion. If <u>naa-</u> is an affix, it is likely to be strictly ordered with respect to other elements, and may `close off' the base to further affixation (Zwicky p. 286). Naa- appears as the leftmost element on the word, but closer to the stem than clitics such as the possessive marker. It cannot be ordered to the right of morphemes such as isht.

(50) nanishtapiisa `ruler'

*ishtnanapiisa

but ã-nanishtapiisa `my ruler'

With this evidence, we can say that naa- appears `closer in' to the base than do clitics, and is thus likelier to be an affix.

Is Naa a Word? While I can provide good evidence than naa is probably not a clitic, there remains the possibility that it is a word. Byington suggests that naa is derived from the word nana `thing; something'. It is possible that naa is a shortened form of nana and that all these derivations are in fact compounds. A number of speakers in fact interchange nana /naa/naa + verb.

(51) nan isht apiisa
thing with measure
`something to measure with'

One interesting fact about naa is that it appears prefixed rather than suffixed, an unusual position for a derivational affix. Some have asserted (particularly Williams (1981) that righthanded headedness, or the placement of the category-specifying morpheme on the right side, is `global', as in his Righthand Head Rule (p. 248): `the head of morphologically complex word is the righthand member.' There are exceptions, even in English: the derivational prefix en- as in endanger, ennoble, enrage would give evidence that the righthand head rule is expected but not mandatory. (Williams acknowledges this counterexample as exceptional, rendering such words headless rather than left-headed.)

Furthermore, there may be reason to be concerned that the syntax is sensitive to the order of the morphological constituents (Baker 1989). Although Baker's case is specific to inflectional affixes, we could make the case that the element that changes the lexical category must be in a position where the syntax could be sensitive to it, or put another way, where its features could project to the phrase level, presumably at word edges. If this matters, we could speculate that Choctaw NPs are relentlessly left-headed, so that in the naa- derivatives the first element of the nominalized verb, and the first element of the NP would be a morpheme signalling `noun'.

We might look at acknowledged compounds to see if the *naa* derivations group with them. Compounds created from two nouns place the `modifying' noun before the `head'.

(52) shokha nipi
pig meat
`pork'

While there are some instances of *naa*-words with bases that are categories other than Verb, these words have unpredictable outputs: cf. (Byington 1915) *naahomma* `red blanketing' (*naa* + A) and even *naabila* (*naa* + N) `grease'. Since the productive *naa* only attaches to verbs, this fact alone would seem to make the compound hypothesis suspect. Why shouldn't we compound *naa* with other `nouns'?

(53) * nipi naa
meat thing
? `meat of unknown origin'

*Naa-* derivations can be compounded like any other noun, however. Szymanek's requirement for compounding is that it involve complete lexical items, suggesting that *naa-* could not appear outside the compound.

(54) naafohka aiitatoba
clothes store
`clothing store'

* naa[fohka aiitatoba]

Running down other possible analyses, since *naa* sits directly to the left of the verb, in the verb's direct object slot, and if not a transitive verb, in the subject slot, we might wonder whether these are not just a verb and its argument.

(55) naa hiɬa
thing dance
`thing dance'

naa haklo
thing hear
`thing hear ~ hear a thing'

But while aknowledging the probable etymology of *naa*-, to accept such a hypothesis means abandoning the validity of both syntax and morphology; it takes little imagination to foresee the linguistic bedlam created by permitting tiny clauses to exist inside noun phrases of this level.

In any event, Choctaw speakers are always able to employ the full word-level noun nana when treating the `thing' idea as a subject or direct object of a verb.

(56) Nana chompa-h sa -banna-h.
something buy pred lacc want pred
`I want to buy something'

From this evidence, I would categorize naa as a nominalizing prefix.

## Conclusion

My analysis of possible morphological nominalizers has been predicated upon two of Aronoff's assumptions; first, that a Word Formation Rule must be `constant and completely specified' (p. 63), the second, that a WFR is independent of the syntax. (In a reformed view, that of Anderson 1990, if morphology is not utterly independent of syntax, at least `all derivation must take place prior to lexical interpretation' (p. 118).)

Given the large number of dictionary-listed nouns appearing in the same form as related verbs, but without predication markers, I first hypothesized that any verb could be nominalized simply by converting it to the desired category. This proved to be incorrect: as only some verbs can be so used, these must be nominalized in the lexicon. These lexicalized nouns do not bear a predictable relationship to the verb from which they are derived; they may be actor, single act of V, condition of V, or others. Again, since the output is not fully specified, this is not a word formation rule.

Missing from the list of output type was gerund: verbs are nominalized in the lexicon at the X-zero level or participate in syntactically nominal constituents.

The second hypothesis was that penultimate accent (high falling pitch) nominalized verbs. This was not upheld experimentally, as native speakers identified verbs as verbs when pronounced in isolation, whether with even stress or penultimate accent, and without predicative -h. A phonemic accent was identified: final accent does form the second person imperative and some aspect markers. Speakers identified final-accented verbs as imperative forms when the verbs were pronounced in isolation. Penultimate accent on verbs may be expected as part of phrasal pitch contour.

Phonemic lengthened vowels in the penultimate may also be associated with higher pitch.

The prefix *naa-* was found to be a successful morphological nominalizer.

## NOTES

*I wish to thank all the Choctaw people who helped me learn about their language: Mae Wilson, Rose Billy, Clelland Billy, Sally Bread, Terry Billy, Evelyn Steele, Dale Steele, Henry Willis, and especially Terry Battiest.

[1] Choctaw orthography has not yet been standardized. Researchers commonly use the popular digraphs <sh> and <ch> for the consonants /š/ and /č/. I follow Ulrich (1986) in his vowel inventory of /a/, /aa/, /ã/; /o/, /oo/, /õ/; /i/, /ii/, /ĩ/. For the lateral fricative, I use the IPA symbol rather than <lh>. Syntactic glossing abbreviations are: acc \`accusative'; nom \`nominative'; pred \`predicate'; subj \`subject marker' ; comp \`complementizer'; dem \`demonstrative'; and asp \`aspect marker'.

[2] We might argue in the case of *ikhaana* \`to know' that this form is the stative/incompletive form of the neutral counterpart *ikhana* \`to learn', denoting an event. Then the \`accent' is obtained by rule and is not underlying.

[3] Ulrich's assertion that glottal stops are diachronically related to other forms may be interesting from another point of view. Robert Ranking (personal communication) suggests that glottal stop in some non-Muskogean languages is often traced to the high-low pitch contour of a parent language (see Rankin 1997). My consultants who are speakers of one dialect in particular, that originating in Louisiana, produce, especially in /oo/ and /ii/ a decided high-low contour in pronouncing some word, such as *tóòbi* contrasting with *tohbi* \`white' and *ishtmíìti* contrasting with *ishtmĩti* \`to bring'.

[4] I use \`nom' for \`n', [?] for ['], and \`pred' for \`v'. Sentence (3) may also be glossed \`I'm looking at the tall man.'

[5] These are the stative, iterative, instantaneous, and intensive, as well as the neutral.

## References

Anderson, Stephen R. 1992. A-Morphous Morphology. Cambridge: Cambridge University Press.

Aronoff, Mark. 1976. Word Formation in Generative Grammar. Cambridge: The MIT Press.

Baker, Mark. 1988. Incorporation: A Theory of Grammatical Function Changing. Chicago: University of Chicago Press.

Booker, Karen. 1988. `The Loss of Preconsonantal *k in Creek/Seminole' IJAL, Vol. 54, No. 4, pp. 371-386.

Broadwell, George. 1990. Extending the Binding Theory: A Muskogean Case Study. Dissertation, University of California at Los Angeles.

Byington, Cyrus. 1870. Grammar of the Choctaw Language. Philadelphia: McCalla and Stavely.

Byington, Cyrus. 1915. A Dictionary of the Choctaw Language. Washington: Government Printing Office.

Chomsky, Noam. 1965. Aspects of the Theory of Syntax. Cambridge, MA: MIT Press.

Chomsky, Noam. 1970. `Remarks on Nominalization' in Jacobs and Rosenbaum (eds) Readings in English Transformational Grammar. Watham, MA: Ginn and Co.

Davies, William. 1986. Choctaw Verb Agreement and Universal Grammar Dordrecht: D. Reidel Publishing.

Downing, Todd. 1974. Chahta Anompa: An Introduction to the Choctaw Language. Durant, OK: Choctaw Bilingual Education Program, Southeastern State College.

Drijkoningen, Frank. 1989. The Syntax of Verbal Affixation. Tubingen, Germany: Max Niemeyer Verlag.

Heath, Jeffrey. 1980. `Choctaw Suppletive Verb and Derivational Morphology' in Kansas Working Papers in Linguistics, Vol. 5, No. 2, pp. 1-24.

Jacob, Betty, Dale Nicklas, and Betty Lou Spencer.

1977. Introduction to Choctaw. Durant, OK: Choctaw Bilingual Education Program.

Klavans, Judith. 1985. `The Independence of Syntax and Phonology in Cliticization' Language, Vol. 61, pp. 95-120.

Lombardi, Linda and John McCarthy. 1991. `Prosodic Circumscription in Choctaw Morphology' Phonology Vol. 8, No. 1, pp.37-72.

Munro, Pamela. 1987. `Some Morphological Difference between Chickasaw and Choctaw' in ed. Munro P. Muskogean Linguistics, UCLA Occasional Papers in Linguistics. #6, pp. 119-133.

Nicklas, Thurston Dale. 1972. The Elements of Choctaw. dissertation, University of Michigan.

Rankin, Robert. 1987. `Ponca, Biloxi, and Hidatsa Glottal Stop and Quapaw Gemination as Historically Related Accentual Phenomena in Mid-America Linguistics Conference 1987 Proceedings.

Sapir, Edward. 1925. `Sound Patterns in Language' in Mandelbaum, David G., ed. Selected Writings of Edward Sapir (1949) Berkeley: University of California Press.

Szymanek, Bogdan. 1989. Introduction to Morphological Analysis Warsaw: Panstwowe Wydawnictwo Naukowe.

Ulrich, Charles. 1986. Choctaw Morphophonology. dissertation, University of California at Los Angeles.

Ulrich, Charles. 1993. `The Glottal Stop in Wester Muskogean' IJAL, Vol.59, No. 4, pp. 430-441.

Williams, Edwin. 1981. `On the Notions "Lexicall Related" and "Head of a Word"' Linguistic Inquiry Vol. 12, No. 2, pp. 245-274.

Wright, Allen. 1860. Chahta Leksikon: A Choctaw in English Definition St. Louis: Presbyterian Publishing Co.

Zwicky, Arnold. 1985. `Clitics and Particles' Language Vol 61, pp. 283-305.

# REFLEXIVES IN MOHAWK

Nancy Bonvillain

Abstract: This paper presents an analysis of meanings and uses of two reflexive morphemes in Mohawk. 'Reflexive' *-atat-* is shown to have both reflexive and reciprocal meanings. It is also realized in kinship terms and in the transitive pronominal prefix *-yutat-* she, one:her, one. 'Semi-reflexive' *-at-* has some reflexive functions. It can also mark middle voice and detransitivized states or processes. Additional uses of *-at-* are exemplified. The paper concludes with discussion of comparative data on reflexives in other Iroquoian languages.

## Introduction

The reflexive morphemes in Mohawk, *-atat-* reflexive and *-at-* semi-reflexive, each have several semantic and grammatical functions. Some of these are clearly reflexive, i.e. indicating actions or processes which affect the subject of the verb in which they are realized. The reflexive *-atat-* also has reciprocal meanings. It occurs as well in a set of kinship terms with a reciprocal sense. Meanings of the semi-reflexive *-at-* are especially complex. This morpheme can mark a number of different kinds of relations between verb bases and their agents or patients. As will be shown, *-at-* can indicate middle voice, detransitivized processes or states, defocusing of agent, spontaneous occurrences or results, and motion or activity away from subject or from prototypical position. Finally, some verbs obligatorily require the selection of a reflexive morpheme as part of a frozen verb stem. Such verbs often have figurative senses.

This paper will examine meanings and uses of the two reflexives in the Akwesasne dialect of Mohawk, one of six extant Northern Iroquoian languages.[1] Following analysis of the Akwesasne material, comparative data from related languages will be discussed. It will be shown that evidence from Mohawk corroborates some of the hypotheses offerred by Barber (1975) and Shibatani (1985) concerning semantic and grammatical relations among reflexives, middle voice and passives. In other respects, though, Mohawk reflexives display a unique constellation of meanings and functions.

Mohawk reflexives occur morphologically within verbs, directly following a pronominal prefix and preceding the verb base which itself is composed of a verb root(s) and an optional preposed incorporated noun root. The structure of Mohawk verbs, then, is:
modal and non-modal prefixes + pronominal prefixes + reflexive or semi-reflexive + noun root + verb root + verbal suffixes.

Only the pronominal prefix and verb root are obligatory categories for all verbs. However, non-imperative verbs (and many imperatives) contain a verbal suffix or suffixes usually marking at least aspect and potentially other kinds of meanings. Most non-habitual verbs contain one of the modal prefixes. Other prefixes express additional meanings.[2] If reflexives are selected, only one of the two is possible, i.e. either *-at-* or *-atat-*. The meaning and uses of these morphemes are discussed and exemplified below. Examples of the reflexive and semi-reflexive are given, with illustration of their contrastive semantic and grammatical functions. The Mohawk data is then discussed in light of other research concerning reflexives, middle voice and related grammatical constructions. Finally, comparative material from other Iroquoian languages is discussed, with a brief note of Caddoan.

Three different sets of pronominal prefixes exist in Mohawk, all of which can co-occur with reflexives although certain verbs constrain the selection of particular classes. The three sets have, in the Iroquoian literature following Lounsbury 1953, traditionally been labelled subjective, objective and transitive. Transitive prefixes express both agent and patient in one form, eg. *-khe-* 'I:her,' *-luwa-* 'she,one:him' as in *khenù:weʔs* 'I like her,' *luwanù:weʔs* 'she likes him.' Subjective prefixes mark the grammatical subject (semantic agent) of a transitive verb while objective prefixes mark the grammatical subject (semantic patient) of an intransitive verb as well as the object of a transitive verb having an inanimate or 'feminine neuter' subject.[3] When the meaning of the reflexive morpheme is narrowly reflexive, i.e. action by agent toward him/herself, then transitive prefixes cannot occur for obvious semantic reasons since they necessarily express a transitive relationship between non-co-referential entities.

As will be shown, some verbs require the reflexive or semi-reflexive exclusively while other verbs may select one or the other, with a significant change in meaning.

### Reflexive *-atat*-[4]

Transitive reflexive meaning: The reflexive *-atat-* most often expresses a transitive action in which agent and patient are co-referential, i.e. action toward self. In such instances, the pronominal prefix co-occurring with *-atat-* must come from the subjective class; subjects therefore are agents. For example:[5]

1a. waʔkátathleneʔ
waʔ-k-atat-hlen-eʔ
fact-I-rfx-cut-asp
I cut myself (eg. with a knife)

1b. Compare non-reflexive verbs:
wahuwáhleneʔ
she cut him (-huwa- she:him)
lahlé:nas
he's cutting (it); he's a cutter (la- he)

Reflexives can occur in transitive verbs which have incorporated their noun root patients:

2a. waʔtkatatesnuhsò:lalakeʔ
waʔ-t-k-atate-snuhs-o:lalak-eʔ
fact-du-I-rfx-hand-press-asp
I squeezed, stepped on my own hand

2b. Compare:
waʔthakesnuhsò:lalakeʔ
he squeezed, stepped on my hand (-hake- he:me)

Reflexives are also used in combination with benefactives in verbs where a transitive action denotes a result for agent:

3. waʔkatathní:nuʔseʔ akyà:tawiʔ
waʔ-k-atat-hni:nu-ʔs-eʔ + akya:tawiʔ
fact-I-rfx-buy-bnf-asp + coat, garment, dress
I bought myself a coat

Finally, reflexives appear in verbs denoting emotional states:

4. latatenù:weʔs
l-atate-nu:weʔ-s
he-rfx-like-asp
he likes himself

Reciprocal meaning: The reflexive *-atat-* has reciprocal meaning in verbs denoting actions of non-singular agents. Since *-atat-* has reflexive meanings as well, its occurrence with non-singular pronominals could potentially be ambiguous. That is, are agents performing action toward themselves or toward each other?[6] Meanings are disambiguated, not by the reflexive which occurs in both cases identically, but rather by inclusion of a dualic verbal prefix -t- (preceding the pronominal) in the composition of reciprocal verbs. The dualic prefix is semantically well-chosen for this function since it expresses the idea that more than one entity participates in the event. For example:

5a. wahyatatkù:lekeʔ
wa-hy-atat-ku:lek-eʔ
fact-they m.du.-rfx-hit-asp
they two hit themselves

5b. waʔthyatatkù:lekeʔ
waʔ-t-hy-atat-ku:lek-eʔ
fact-du-they m.du.-rfx-hit-asp
they two hit each other

5c. waʔkyakyatátkwahteʔ
waʔ-k-yaky-atat-kwaht-eʔ
fact-du-we du.excl.-rfx-bump-asp
we two bumped into each other

Reciprocal usage in 6a. below contrasts with a non-reflexive form in 6b:

6a. wakatsenú:niʔ ǰiʔ tusaʔkyatátkʌʔ
wak-atsenu:niʔ tu-s-a -ky-atat-kʌ-ʔ
I-be happy-asp that du-rep-fact-we du.incl.-rfx-see-asp
I'm glad that we two saw one another again, I'm glad that we two met again

6b. wakatsenú:niʔ ǰiʔ saʔkú:kʌ
I'm glad to see you again (-ku- I:you)

If the verb base *-atatkʌ-*, containing the reflexive, occurs with a non-singular agent but without the dualic, then a reflexive meaning is conveyed:

6c. ká:ts yekyatátkʌʔ
ye-ky-atat-kʌ-ʔ
trns-we du. incl.-rfx-see-asp
Come, let's look in the mirror!

Since this verb lacks the dualic, it must mean 'let's look in the mirror (at ourselves)' rather than 'let's look at each other.'

Figurative uses: The reflexive *-atat-* often occurs in verbs which have taken on figurative meaning, denoting events which are not literally transitive. In some cases, a reflexive meaning is still semantically isolable and obvious whereas in others it has become submerged as part of a verb with a unified figurative sense. Several examples follow, contrasted with non-reflexive verbs:

7a. skatateʔnikuhlʌhtʌ́:nis
s-k-atate-ʔnikuhl-ʌ-ht-ʌ:ni-s
rep-I-rfx-mind-happen-cs-bnf-asp
I make myself forget (I cause to happen to my mind)

7b. sewakeʔnikúhlʌ
I forgot (-wake- I)

8a. tewatatelihwanelakhseláhlos
tew-atate-lihwa-nelak-hsela-hl-o-s
we pl.incl.-rfx-idea-bad-nom-coat-be,put in water-asp
we all stain ourselves with sin (we put a liquid coating of bad ideas on ourselves)

8b. wahsatatkuhtseláhluʔ
wa-hs-atat-kuht-sela-hl-uʔ
fact-you-rfx-paint-nom-coat-asp
you got paint on yourself

8c. láhlos okúhtselaʔ
he's painting (it) (he's coating + paint)

9a. yaʔkatatewʌní:yohsteʔ
y-aʔ-k-atate-wʌn-i:yo-hst-eʔ
trns-fact-I-rfx-word,sound-be good,nice-cs-asp
I helped myself over there (eg. took something from my neighbor's house)(I caused, made the word good)

9b. lowʌniyóhstuʔ
he gave his full consent (-lo- he)

10a. yakotateʔnikuhlotákwʌ
yako-tate- ʔnikuhl-ot-akw-ʌ
she-rfx-mind-put in-rvrs-asp
she is tempted (she takes her mind out, it

takes her mind out)

10b. liʔnikuhlotákwas
I tempt him (-li- I:him)

11a. sahatathlé:wahteʔ
s-a-h-atat-hle:waht-eʔ
rep-fact-he-rfx-punish-asp
he's sorry for it, sorry about it (he punished himself)

11b. í:kehleʔ   ausakatathlé:wahteʔ
i:-k-ehl-eʔ   au-sa-k-atat-hle:waht-eʔ
**-I-want-asp   opt-rep-I-rfx-punish-asp
I want to take out a pledge, make a pledge (eg. to abstain from wrong behavior)

11c. sasatathlé:waht
sa-s-atat-hle:waht
rep-you-rfx-punish
repent! ask for forgiveness! (punish yourself)

11d. í:kehleʔ   ausakoyatathlewáhtʌ
i:-k-ehl-eʔ   au-sa-koy-atat-hlewaht-ʌ
**-I-want-asp   opt-rep-I:you-rfx-punish-asp
I want to ask your forgiveness

In 11d., *-atat-* co-occurs with a transitive pronominal, -*koy-* 'I:you.' The reflexive is not referenced semantically to the entire pronominal but only to the agent 'I.' The verbal construction, then, has an underlying meaning approximating 'I want to ask you that I punish myself.'

Kinship terms: The reflexive *-atat-* appears in the formation of several kinship terms. All Mohawk kin terms are structurally verbs, each containing a pronominal prefix, verb root and aspect suffix. The verb root denotes an underlying consanguineal or affinal meaning while the pronominal prefix indicates a specific status relationship between speaker and referent (or addressee). Most terms select transitive prefixes marking agent/patient co-relations. The elder person in the kin pair (whether speaker or referent) is expressed as agent while the younger is patient. For instance, *kheʔkʌ́:ʔaʔ* 'my younger sister' (*khe-* I:her), *laksóthaʔ* 'my grandfather' (*lak-* he:me).

Another set of kin terms selects subjective + reflexive pronominals. These express equality and/or unity of the named kinsperson and the speaker. The subjective pronominal is first person dual, *yaky-*, when reference is to one relative (speaker + relative) and is plural, *yakw-*, when two or more relatives are named (speaker + relatives). The morpheme *-atat-* marks the co-relationship as reciprocal. Such kin terms include:

12a. yakyatatenúhkweˀ
yaky-atate-nuhkw-eˀ
we du.-rfx-be kin-asp
we two are related, we're relatives, s/he is
my relative, my relative

12b. yakwatateˀkʌ́:ˀa
yakw-atate-ˀkʌ:-ˀa
we pl.-rfx-have a younger sibling-asp
my siblings, we are all siblings

12c. yakyatakyʌ́:ˀa
yaky-atak-yʌ:-ˀa
we du.-rfx-have as child-asp
me and my daughter or son or mother or father
(any group of two of us, parent and child)

Similar kin terms can be constructed with third person dual or plural subjective pronominals + reflexive in reference to relatives of another person. Among these are:

12d. kyatateˀkʌ́:ˀa
ky-atate-ˀkʌ:-ˀa
they f.du.-rfx-have a younger sibling-asp
her sister

12e. lutakyʌˀsú:ˀa
lu-tak-yʌ-ˀsu:ˀa
they m.pl.-rfx-have as child-many,diverse
they are parent(s) and child(ren) (group of
three or more)

Finally, although most kin terms have the same basic construction in reference or address (except for change in the aspect suffix, eg. *laksótha*ˀ 'my grandfather,' *lákso* 'grandfather!'), a revealing optional form exists for sibling address terms. Instead of the transitive pronominals khe- 'I:her' or li- 'I:him,' the subjective prefix expressing first person dual inclusive + reflexive is possible:

12f. shé:ku kyatatè:kʌˀ
ky-atate:-kʌ-ˀ

we du.incl.-rfx-have as younger
sibling-asp
hello sister! hello brother!

This usage has an obvious pragmatic rationale, signifying, as it does, a reciprocal social relationship between speaker and addressee.

Transitive pronominal -*yutat*-: The reflexive -*atat*- occurs in the composition of the transitive pronominal prefix -*yutat*- which expresses feminine indefinite agent 'she,one' and a non-identical feminine indefinite patient, 'her,one' (see note 3). This pronominal consists of the subjective marker -*ye*- she,one + reflexive -*atat*- and undergoes regular Mohawk morphophonemic changes whereby ye ⟶ yu/____a with subsequent loss of -a-. For example, *yutatenu:we s* 'she likes her,' *yutateku:leke* 'she will hit her.'

These words, and others similarly formed, are ambiguous between transitive and reflexive meanings since the transitive pronominal is internally composed of subjective -ye- + reflexive -atat-. For instance, *ʌyutatekù:lekeʔ* can mean 'she will hit her' or 'she will hit herself.' The two senses can be disambiguated contextually or lexically for a transitive meaning by specifying referents. In such constructions, agents precede verbs while patients follow:

13a. katelí ʌyutatkù:lekeʔ neʔ sesí
Catherine will hit Cecelia

13b. katelí ʌyutatkù:lekeʔ né:ʔe
Catherine will hit her

13c. akáuha ʌyutatkù:lekeʔ né:ʔe
she's the one who will hit her

Reflexive meaning cannot be specified lexically since Mohawk does not have a separate word for self. However, if only an agent is realized, as in *kateli ʌyutatkù:lekeʔ* , the construction is given a reflexive rather than transitive reading, i.e. 'Catherine will hit herself.'

Semi-reflexive -*at*-:

Semi-reflexive -*at*- has a number of meanings and usages. The semantic connection among its various senses is not easily summarized. Some occurrences have a reflexive meaning in which agent (if subjective) or patient (if objective) is instigator or experiencer. Other uses of *-at-* indicate a kind of middle voice. One such use encodes the "involvement" (Barber 1975:17) of subject in the action. In these verbs, the subject "performs the action but also receives some benefit" or is in some other way affected by their action (ibid:19). Another kind of middle voice signalled in Mohawk by -at- occurs in verbs in which an agent is assumed but not explicitly marked. In these verbs, agents are defocused and ultimately suppressed, raising patients to subject. In such detransitivized verbs, objective pronominals (expressing patient subjects) are employed.

In addition to these functions of -at- which relate in some way to voice, -at- can occur in verbs denoting a spontaneous occurrence or state. Another use of -at- is to mark activity or movement away from agent or prototypical position. Finally, some verbs require inclusion of -at- as part of a frozen verb base. These verbs never occur without -at- and therefore are not contrastive. Throughout this section, contrastive examples will be given where possible.

One usage of *-at-* signifies a reflexive event in which an agent acts toward him/herself. These verbs have transitive bases, although some have been detransitivized and function as intransitives. Figurative senses can be produced by this use of *-at-*:

14a. wahatáhsehteˀ
wa-h-at-ahseht-eˀ
fact-he-srfx-hide-asp
he hid himself

14b. waˀúhsehteˀ
she hid it (-u- she,one)

15a. waˀthyateˀkháhshiˀ
waˀ-t-hy-ate-ˀkhah-shi-ˀ
fact-du-they m.du.-srfx-border-undo-asp
they two separated, divorced

15b. waˀtkeˀkháhshiˀ
I took it apart, separated it (into two pieces; -ke- I)

| | | |
|---|---|---|
| 16a. | isenekwá | sathá:wiht |
| | ise-nekwa | s-at-ha:wiht |
| | there-direction | you sg.-srfx-slide,move |
| | slide over there! | move over there! |

16b. kʌhnewká waˀehá:wihteˀ
she moved, slid it over here (without picking it up; -e- she,one)

17a. wahatelákwʌ
wa-h-ate-lakw-ʌ
fact-he-srfx-choose-asp
he kept it (chose it for himself)

17b. with incorporated noun root, -wil- 'baby' 'adopt':
waˀkatewilalákwʌ
I adopted her (chose baby for myself)

17c. wahalákwʌ
he chose it (-ha- he)

Closely associated semantically with a direct reflexive meaning, the semi-reflexive can be used to indicate greater involvement, energy-expenditure or concentration of agent in an event as compared to non-reflexive verbs. In these examples, subjects accomplish an act by themselves or accomplish an act which absorbs their energy:

18a. yuthnawelú:niˀ
yu-t-hnawel-u:ni-ˀ
she,one-srfx-well-make,build-asp
she's digging, making a well (by herself)

18b. yehnawelú:niˀ
she's digging a well (someone, unspecified is helping her; ye- she,one)

19a. tesatohtáhlos kʌˀ
te-s-at-ohtahlo-s
du-you sg.-srfx-clean-asp + int
are you cleaning up? (implies a big job)

19b. tesohtáhlos kʌˀ
are you tidying up? (-s- you)

The following sentence, using the verb *-ohale-* 'wash' is directly contrastive for the semi-reflexive:
20. waˀkatenuhsóhaleˀ náhkweˀ thonó:nʌ
waˀ-k-ate-nuhs-ohale-ˀ first then
fact-I-srfx-house-clean-asp

waˀkeksohalényuˀ
waˀ-ke-ks-ohale-nyu-ˀ
fact-I-dish-clean-dst-asp

I washed the floor first, then I washed the dishes

Here, washing a floor inherently requires more energy, time, etc. than washing dishes and therefore selects the semi-reflexive.

An important use of -at- marks a kind of detransitivized event or state derived from a verb whose basic non-reflexive meaning is a transitive action. These are middle voice or passive constructions. They involve a "decrease in semantic valency [of an agent], to the extent that no agent is posited" (Shibatani 1985:839). In Mohawk, morphologically or syntactically-marked passive constructions do not occur. Instead, one function of the semi-reflexive is to express middle voice or detransitivized meanings. Therefore, some shifts in transitivity or voice are realized through use of -at-

The semi-reflexive co-occurs with animate patients (in objective case) to indicate a passive experiencer of an event or process. Agents are suppressed. Figurative senses can be produced by such use of -at-. The following examples contrast verbs containing -at- and non-reflexive transitives employing the same verb roots:

21a. waʔtewakatenuhyaníhtʌ
wa ʔ-te-wak-ate-nuhyaniht-ʌ
fact-du-I-srfx-exaggerate,be extreme-asp
I was shocked (it was exaggerated, extreme to me)

21b. tehanuhyaníhtaʔ
he's an exaggerator (-ha- he)

22a. yakoteǰilʌ́htuʔ
yako-te-jil-ʌ-ht-uʔ
she,one-srfx-burning cinders-fall,drop down-cs-asp
she is damned in hell, she's gone down into the fire
22b. yaʔkǰí:lʌhte ʔ
I made burning cinders drop down (eg. into a stove; -k- I)

23a. wakateʔnikuhliyóhstuʔ
wak-ate-ʔnikuhl-iyo-hst-uʔ
I-srfx-mind-be nice,good-cs-asp
I'm being patient (I'm making my mind nice, good)

23b. wʌhniseliyóhsteʔ
she makes the day nice (w- she,it; -ʌhnisel- day)

The verb base in 23a. *-ʔnikuhliyo-* 'be nice-minded,' can contrast with a non-reflexive, non-causative construction. When the base co-occurs with -at- + causative (23a.), it describes a specific event entailing an on-going state which may or may not have preceded or be expected to endure. However, when both affixes are lacking, eg. *wakeʔnikuhlí:yoʔ* 'I'm patient (my mind is nice, good)' the verb denotes a generalized state of being or an inherent characteristic.

When -at- co-occurs with inanimate patients as grammatical subjects (marked by objective pronominals), verbs denote results of processes. Agents are suppressed; patients fulfill semantic roles as experiencers. In these cases, verbs are either in punctual aspect, indicating a single event, or in perfective state aspect, indicating a generalized resultant state:

24a. yeyotelò:luʔ
ye-yo-te-lo:l-uʔ
trns-she,it-srfx-gather-asp
it's gathered, accumulated there

24b. selò:lok
gather it! (se- you)

25a. ʌsewatelíhshiʔ
ʌ-se-w-ate-lihsh-iʔ
fut-rep-she,it-srfx-undo-asp
it will be undone

25b. ʌsewatuhwʌǰalíhshiʔ
the world will become undone (-at- srfx; -uhwʌǰa- 'world, earth')

25c. sahalíhshiʔ
he undid it (-ha- he)

The semi-reflexive can be employed to mark an activity which occurs spontaneously or a state which results from unspecified, spontaneous conditions.[7]

26a. waʔkutehyá:luʔ
waʔ-ku-t-ehya:l-uʔ
fact-they f.pl.-srfx-grow, raise-asp
they all grew up

26b. sehyá:luʔ
make it grow! raise it! (s- you)

26c. í:ʔi waʔkheyehá:luʔ
I raised her, I'm the one who raised her
(i:ʔi I; -khey- I:her)

27a. teyokyà:kuʔ
te-yo-k-ya:k-uʔ
du-she,it-srfx-break in two pieces-asp
it's broken (into two pieces)

27b. úʔkyaʔkeʔ
uʔ-k-yaʔk-eʔ
fact-she,it-srfx-break-asp
it broke

The verb in 27a. denotes breaking but does not specify an agent. It implies that an event occurred without an immediately direct agent although some agent may be responsible for starting the process of breaking. This word could be used, for instance, to describe the breaking off of a pole in the ground which rots and finally breaks at ground level (into two pieces).

The verb root *-yaʔk-* 'break' can co-occur with incorporated noun roots in either semi-reflexive or non-reflexive constructions indicating a contrast between a spontaneous result without stated agent (semi-reflexive) and a direct action with explicit agent (non-reflexive). Compare the following verbs having animate subjects:

27c. wahanitáhsyaʔkeʔ
wa-h-an-itahs-yaʔk-eʔ
fact-he-srfx-tail-break-asp
his tail broke off, fell off

27d. wà:kehleʔ waʔtkačínyaʔkeʔ
I thought I broke my leg (-ka- I; -čin- leg)

A further use of *-at-* is attested in verbs denoting change in body position away from the prototypical position of upright posture, eg. sit down, kneel down. In these cases, the semi-reflexive verb emphasizes the process of changing position while the non-reflexive verb signifies the resulting state:

28a. waʔtkatʌtsó:tʌ
waʔ-t-k-at-ʌtso:t-ʌ
fact-du-I-srfx-kneel-asp
I knelt down

28b. teyakʌ́tskoteʔ

she's kneeling down (-yak- she,one)

29a. sanitskó:tʌ
s-an-itsko:t-ʌ
you sg.-srfx-sit-asp
sit down!

29b. kítskoteʔ séʔ
I am sitting down (k- I; seʔ emphatic)

Another use of -at- which may be related semantically to that shown in 28 and 29 occurs in the differentiation of verbs meaning buy/sell and borrow/lend. In these sets, -at- marks movement of goods away from agent (sell, lend) while the non-reflexive verb denotes movement toward agent (buy, borrow). Compare:

30a. waʔkatʌhní:nuʔ
waʔ-k-atʌ-hni:nu-ʔ
fact-I-srfx-buy-asp
I sold it

30b. waʔehní:nuʔ
she bought it (-e- she,one)

31a. wahatʌníhaʔ
wa-h-atʌ-ni-haʔ
fact-he-srfx-borrow-asp
he lent it out

31b. wahaníhaʔ
he borrowed it (-ha- he)

The semantic similarity between these examples and those in 28 and 29 is suggested by the fact that in both sets the semi-reflexive marks verbs which encode "movement away", i.e. movement away from agent (sell, lend) or away from prototypical human position (sit down, kneel down).[8]

In constructing some transitive verbs, realization of -at- is correlated with incorporation of a noun root indicating direct object. These usages can be literal or figurative:

32a. waʔkatenaʔtalú:tʌ
waʔ-k-ate-naʔtal-u:t-ʌ
fact-I-srfx-bread-put in, attach-asp
I baked bread (put bread inside)

32b. waʔkú:tʌ

I attached it, put it inside (-k- I)

33a. laniǰató:lats
l-an-iǰ-ato:lat-s
he-srfx-fish-hunt-asp
he's a fisher, he fishes

33b. latelihwató:lats
l-ate-lihw-ato:lats
he-srfx-idea-hunt-asp
he's a gossip, he gossips

33c. yutó:lats
she's a hunter, she hunts (yu- she,one)

34a. kwah ukwatelihwáhtuˀseˀ ǰiˀ nahó:tʌ waˀkyóhteˀ
ukw-ate-lihw-ahtu-ˀs-eˀ waˀ-k-yoht-eˀ
very fact+I-srfx-idea-lose-bnf-asp what fact-I-work-asp
I lost interest in what I was doing

34b. ukwáhtuˀseˀ
I lost it, it disappeared to me (ukw- fact+I)

Reflexive/semi-reflexive contrasts

Some verbs can select either *-atat-* or *-at-*. Significant semantic changes in the unitary meaning of words result from realization of each morpheme. Several comparisons follow. In the first, verbs entail differences in passive or active role of subjects. In 35a., with semi-reflexive -at-, the subject undergoes a process while in 35b., with reflexive -atat-, she initiates the process which affects her reflexively. This contrast highlights use of -at- in a middle voice construction while -atat-is employed for reflexive meaning. Furthermore, the semi-reflexive verb signals less control on the part of subject or experiencer (see Jelinek and Demers 1983 for discussion of middle-passives and lack of control.)

35a. wakathkalú:niˀ
wak-at-hkal-u:ni-ˀ
I-srfx-suffer-make-asp
I had a misfortune, a misfortune happened to me

35b. yakotathkaluniháhkyeˀ
yako-tat-hkal-uni-hakye-ˀ
she,one-rfx-suffer-make-prog-asp
she continually, repeatedly brings misfortunes

on herself

Another example of contrast between active, voluntary instigation, marked by *-atat-*, and unspecified, uncontrolled endurance of a process, marked by *-at-*, involves the verb root *-uni-* 'make, build.' With neither *reflexive* prefix, the verb denotes a transitive action, eg. *hlu:ni* 'he'll build it' (-hl- he). The verb with -at- (36a.) denotes a spontaneous event or state with suppression of agent while -atat- (36b.) indicates controlled, deliberate activity:

36a. katú:niʔ
k-at-u:ni-ʔ
I-srfx-make-asp
I am born

36b. kwískwis yakotatú:niʔ
yako-tat-u:ni-ʔ
pig she,one-rfx-make-asp
she made herself into a pig

This sentence has a number of associated usages: she dressed up like a pig; she turned herself into a pig (she's a witch and transforms her bodily form); she eats like a pig. The verb in 36b. can further contrast with a semi-reflexive form, highlighting the significant difference in activity. In 36c., the agent is defocussed:

36c. kwískwis lotú:ʔu
lo-t-u:-ʔu
pig he-srfx-make-asp
he became a pig, turned into a pig, was made into a pig

In both 36b. and 36c., the subject is expressed as patient (with objective pronominals), indicating that s/he experiences the result of an action or process. However, in 36b., with -atat-, the patient voluntarily undergoes the process, in fact causing it to happen, whereas in 36c., with -at-, the patient is a passive experiencer.

In the next contrasting set, differences among the verbs seem to entail more focussed concentration and self-involvement in reflexives (37c., 37d.) as compared to non-reflexives (37a., 37b.) and a deeper internalization and affective state in the reflexive (37d.) as compared to a semi-reflexive (37c.):

37a. seʔnikù:lalak
se-ʔniku:l-al-ak
you sg.-mind-watch,guard-cont
be careful!

37b. tehiʔnikúhlaleʔ
te-hi-ʔnikuhl-al-eʔ
du-I:him-mind-watch,guard-asp
I'm bothering, disturbing, worrying, pestering him

37c. wakateʔnikù:lalʌ
wak-ate-ʔniku:l-al-ʌ
I-srfx-mind-watch,guard-asp
I was careful about it, guarded it, watched out for it

37d. tewakatateʔnikúhlaleʔ
te-wak-atate-ʔnikuhl-al-eʔ
du-I-rfx-mind- atch,guard-asp
I'm worrying, I worried about it.

A final contrast presents verbs containing the base for 'lock,' *-nyʌthahlok-*. Without reflexive morphemes, the verb has a simple transitive sense, eg. *ʌhanyʌtháhlokeʔ* 'he'll lock it.' With -at- and an inanimate patient as subject (38a.), the verb denotes a spontaneous, uncontrolled event in which the agent is defocussed. It refers to self-locking, as when a lock repeatedly gets stuck or is locked without animate intervention:

38a. watenyʌtháhloks
w-ate-nyʌthahlok-s
it,she-srfx-lock-asp
it keeps locking itself

Next, compare the following two verbs with animate agents. The first, 38b., contains -at-; 38c. contains -atat-:

38b. waʔkatenyʌtháhlokeʔ
waʔ-k-ate-nyʌthahlok-eʔ
fact-I-srfx-lock-asp
I locked myself in

38c. waʔkatatenyʌtháhlokeʔ
waʔ-k-atate-nyʌthahlok-eʔ
fact-I-rfx-lock-asp
I locked myself out

The contrast here seems to differ from the distinction shown in verbs for buy/sell, borrow/lend and sit down, kneel down (see examples 28 through 31). In those cases, the semi-reflexive contrasted with non-reflexives to indicate movement away from subject (sell/lend) or prototypical position (sit down, kneel down). In 38b., 'I locked myself in,' -at- indicates inward activity or possibly activity toward prototypical place (inside, eg. inside one's house) whereas in 38c., 'I locked myself out,' -atat-, which is a reduplication of -at-, indicates outward activity or activity away from prototypical place.

Discussion

Evidence from Mohawk suggests an association among various functions of reflexives and the creation of middle voice, detransitives and defocusing of agents. Shibatani (1985) demonstrates similar interconnections among these grammatical constructions in a wide sampling of languages from Europe, Asia and the Americas. Shibatani suggests that reflexives and reciprocals are fundamentally related to passives, the prototype of which entails defocusing of agents (1985:837). Since there are often "...semantic and syntactic similarities among passives, reflexives, and reciprocals, for some languages, these similarities are sufficient to permit one form to perform the functions of the others" (ibid:843). In Mohawk, in the absence of distinctive passive formations, reflexive morphemes are sometimes employed to indicate passive meanings, both with animate subjects (eg. 'I was shocked') and inanimates ('it's gathered there').

Shibatani further correlates reflexives with middle-passives which express spontaneous events, denoting activities or states which occur automatically or without intervention of agents (ibid:827). Mohawk -at- fulfills this function as well, eg. 'they all grew up.'

Some Mohawk verbs with semi-reflexives indicate states resulting from actions or processes. Agents may be implied but are not explicitly encoded, eg. 'it broke into two pieces.' Other verbs can have stative surface meanings, eg. 'I'm being patient.' Both Fagan (1988:195) and Keyser and Roeper (1984:383) claim that middle verbs in English are essentially of this type, i.e. statives. Although Mohawk semi-reflexives can signal stative meanings, the morphemes have a wider range of functions. Furthermore, many surface statives such as 'I'm being patient' have underlying transitive bases ('I'm making my mind nice').

In a study of Koyukon Athapaskan, Thompson (1989) draws a connection among reflexives, reciprocals and passives (and causatives), based on the fact that they are "...involved in constructions related to transitivity" (1989:9). In Mohawk, reflexives are employed to signify transitive actions toward self or to mark detransitivity. Thus, verbs which have basic transitive meanings can be rendered intransitive (eg. 'slide, move over there!').

Finally, Jelinek and Demers (1983) suggest a use of middle voice to mark "...less control on the part of the subject of the sentence" in some Coast Salish languages (1983:182). Some Mohawk examples attest to this function of semi-reflexives as well (eg. 'I had a misfortune'). Such constructions contrast with the responsibility and control signalled by reflexives ('she continually brings misfortune on herself').

For European languages, Barber (1975) describes middle voice as a "...strategy for marking identities between the surface subject and other NPs in the sentence" (1975:17). In middle voice, non-agentive NPs are raised and merged with subjects. Reflexives in Mohawk perform this function in their role to mark actions by subjects which affect themselves or each other. They also mark an action which subject performs for his/her benefit. Indeed, one of Barber's Greek examples has an exact parallel in Mohawk. Note the following (ibid:18):

| | |
|---|---|
| hair-oumai | moiran |
| take-mid.voice | share |

I choose (take for my own benefit) a share

In Mohawk, *wa-h-ate-lakw-* 'he kept it (chose it for himself)' is formed on the same principles.

Barber points out that middle voice, reflexive, reciprocal and passive constructions all share a semantic feature which distinguishes them from true active voice, i.e. they all refer to events in which subjects are somehow affected by their actions (ibid:20). In true actives, in contrast, subjects are not affected by their actions.

However, the ways in which reflexives and other constructions affect their surface subjects differ. In reflexive formations, subjects affect themselves; in reciprocals, subjects affect each other; and in middle voice and passives, subjects are affected by external agents which may or may not be explicitly mentioned. Some of these distinctions are noted in Mohawk. For example, reciprocal meanings are conveyed by -atat-; reflexive meanings are signalled by either -atat- or -at-, both co-occurring with subjective pronominals; middle voice and passive meanings are marked by -at- co-occurring with objective pronominals.

The role of -at- as an indicator of voice is a demonstration of Hopper and Thompson's (1980) theses concerning degrees of transitivity. As they explain, reflexives have "...an intermediate status between one-argument and two-argument clauses" (1980:277). Reflexives are higher in Transitivity than prototypical one-argument clauses (intransitives) yet lower in Transitivity than prototypical two-argument clauses (true transitives). As the data above has shown, the semi-reflexive in Mohawk performs functions which decrease the transitivity of many verbs (eg. I had a misfortune; slide over there!).[9]

Barber suggests that in Indo-European languages, reflexives, reciprocals and passives developed diachronically from middle voice. In these languages, such as English, reflexives and reciprocals arose through the development of specialized pronouns. Middle voice in verbs has become realized through passives (1975:21-22). It may well be that in Mohawk, the reflexive -atat- is a reduplication of and historically derived from the semi-reflexive -at-, the latter morpheme having a productive function in middle voice constructions.

Comparative Evidence

Northern Iroquoian: Mohawk reflexives -atat- and -at- have cognates in all Northern Iroquoian languages. Underlying phonological shapes are identical although some allomorphic rules diverge.[10] In languages most closely related to Mohawk, i.e. the Five Nations of Oneida, Onondaga, Seneca and Cayuga, basic reflexive forms and grammatical functions are the same (see Lounsbury 1953:72-75 for Oneida; Chafe 1967:26-27 for Seneca; Chafe 1970:48-49 and Woodbury 1975 for Onondaga). Lounsbury estimates that these Five Nations languages separated historically from the other Northern branch of Tuscarora/Nottoway approximately two thousand years ago (1978:335). For Tuscarora, Williams (Mithun) describes one reflexive, -at-, which has both reflexive and reciprocal senses, eg. *yakya:tkʌh* 'we two see ourselves' or 'we see each other' (1976:160). Evidently, the dualic prefix, which does exist in Tuscarora, is not used to disambiguate these meanings as it is in Mohawk and other Northern Iroquoian languages. Rudes (1981) exemplifies two reflexives in Tuscarora and the closely-related but now extinct Nottoway, earlier spoken in Virginia and known through word lists collected by John Wood in 1820 (1981:27). Allomorphic rules in Tuscarora and Nottoway differ from those of the other Northern Iroquoian languages, a result of overall sound changes. From Rudes' data, -atat- has both reflexive and reciprocal meanings while -at- has reflexive and other functions similar to those operating in Mohawk. Among Rudes' examples are: 'it struck itself' in T: ə̃ʔ*nat* *kə̃* *hruk* and in N: *untatenheerug*; 'your own ears' in T: θ*athə̃hnəh* ~~h~~ and '(you) listen' in N: *satuntatag* (ibid:36, 40).

Two additional Northern Iroquoian languages no longer spoken, Huron and Wyandot, also had reflexive and semi-reflexive morphemes, evidently with forms and functions similar to those in present-day Mohawk. Among Barbeau's examples from Wyandot, spoken until recently by Oklahoma descendants of aboriginal Huronian peoples of Ontario, are: *a*ʔ*-y-at-rih* 'I-self-office hold' and *h-(a)tate-n-ate:-ya*ʔ 'he-to self-hair burns' (1915:8, 13).

Another extinct Northern language, currently known as Laurentian (Lounsbury 1978:335), has some scanty sixteenth-century documentation added to journals of Jacques Cartier. J.A. Cuoq cites the following words from Cartier's lists and from nineteenth century Mohawk: 'look at me' (regarde-moi) in M: *takatkatho* and Laurentian: *quatgathoma* (1882:188). This example attests to realization of -at- in the latter language.

The fact that reflexive functions are old in Mohawk is shown by another work of Cuoq. In discussing reflexives in nineteenth century Mohawk, he notes that -atat- had both reflexive and reciprocal meanings and that the dualic -te- was used in reciprocals with non-singular subjects (1866:111). According to Cuoq, -at- functioned to render verbs passive but also occurred in deponent verbs ("verbes deponents") which have passive form but active meanings (ibid:112-113).

Cherokee: Southern Iroquoian is represented by Cherokee, spoken in North Carolina and also by descendants of the aboriginal population who now reside in Oklahoma. Lounsbury estimates branching of Cherokee from the Northern languages to have occurred approximately 3500-

4000 years ago (1978:334). Like Northern languages, Cherokee has reflexive and reciprocal markers. Reyburn describes a reflexive -adad-″ which is positioned following a pronominal prefix and preceding a verb stem (1953:267). From his examples, -adad- seems to have reflexive meanings when the subject is singular and reciprocal meanings when the subject is dual or plural. For instance, *zadadeyohəsga* 'I am teaching myself'; *doznadadeyohəsga* 'another and I are teaching each other;' and *dandadeyohəsga* 'they teach each other' (ibid:268, 269, 272). In these examples, verbs with non-singular subjects contain a prefix d- which Reyburn earlier defines as 'plural' (ibid:178) and which is cognate to Mohawk dualic prefix -t- co-occurring with non-singular subjects + reflexive -atat- to render a reciprocal rather than reflexive meaning. Reyburn does not cite examples of -adad- and dual or plural subjects without the prefix -d-. Since he does not give any attention to this issue, it is not possible to know whether such co-occurrences may or may not exist, or to know how one would express reflexive meanings for non-singular subjects.

Reyburn's material reveals another form which has some reflexive as well as other functions. It is realized as -al-, -ali-, and -ał-. One contrasting pair is given: *zisdoyeha* 'I am shaving him' and *galisdoyeha* 'I am shaving myself' (ibid:273). Other examples of this prefix are: *gali:tuwoá* 'I am combing my hair;' *gali:eligá* 'I am happy;' and *hałsgisgoí* 'you habitually dance' (ibid:175, 177).

In an analysis of Cherokee text material, Walker segments the prefix ali- (ał/__s) which he translates as 'with reference to oneself' (1975:214). It occurs in numerous examples with reflexive meanings including: *galisde:liha* 'I'm helping myself' (compare: *dejisde:liha* 'I'm helping them'); *agwałsdayɨ́hɨsga* 'I'm being eaten' (something is feeding itself on me); and *hałsdɨ:tluga* 'sit down!' (ibid:223-224).

These Cherokee materials show clear resemblances to Mohawk reflexives, encompassing similarity in phonological form, structural position within verbs, and overlapping of semantic functions. The data obviously attest to the antiquity of reflexive marking in Iroquoian.

Caddoan: Since Chafe (1976, 1979) has suggested a historical connection between Iroquoian and Caddoan, it is relevant to turn briefly to Caddoan languages for evidence of reflexive forms. In Chafe's discussion of Caddoan and Iroquoian relationships, he illustrates a correspondence between the Seneca semi-reflexive -at- and Caddo benefactive -t-.

In addition to phonological resemblances, Seneca semi-reflexive and Caddo benefactive share structural similarities since both occur as "leftmost constituents of verb bases" (1976:52). Both follow pronominal prefixes. Chafe further draws a parallel in their semantic and grammatical functions, i.e. they "affect transitivity in some of their occurrences, although in opposite ways. The Caddo 'benefactive' sometimes transitivizes bases otherwise intransitive, while the Seneca element sometimes has the opposite effect"(ibid) (see Mohawk examples above).

Among Chafe's Caddo examples of benefactives (-t- is sometimes realized as -n-, an alternation attested in Mohawk reflexives preceding the vowel -i-): *tʔu-t̲-háy-yúh-cáh* 'I am going to tell to you;' *ku -'n̲-kiyuʔ-ah* '(he) killed for me;' and *ku-n̲-biʔn-čah* '(he) is going to wipe for me' (ibid:73). Chafe further states that Caddoan benefactive also has a possessive sense.[12] For instance, *hák-ku-t̲-cakah-ʔaʔ* 'a belt is present to me, I have a belt' (ibid: 74). The Mohawk example, *wahatelákwʌ* 'he kept it (he chose it for himself)' provides evidence of a benefactive function of the semi-reflexive.

In his review of the Caddoan family, Chafe (1979) reconstructs proto-Caddoan benefactive as -t- or -r-. It occurs in Northern Caddoan as -ut- or -ur- (eg. Wichita, see Rood 1976). Chafe suggests that the -u- was "originally part of an agreement prefix" and links this to Iroquoian semi-reflexive -at- (-an- or -ar-) which "may also have picked up its vowel from preceding agreement prefixes" (1979:230-231).

In sum, then, comparative evidence indicates possible connections among Iroquoian and Caddoan morphemes which share similarities, although not identities, in phonological shape, structural position within verbs and in semantic and grammatical functions. Although data are not definitive, they are suggestive and supportive of hypotheses concerning familial relationships. Increasing research in comparative reconstructions of grammatical and semantic categories will hopefully provide confirmation of these hypotheses in the future.

# NOTES

[1] The data examined come from Akwesasne Mohawk, spoken at the Akwesasne (St. Regis) Reserve, located in northern New York State and in Canada in Ontario and Quebec. The Reserve has a total population of approximately 10,000, possibly half of whom are speakers of Mohawk with varying degrees of proficiency.

[2] For discussion of the structure of Mohawk verbs and other details of grammatical and phonological features of the language, see Bonvillain 1972, A Grammar of Akwesasne Mohawk, and Beatty 1974, Mohawk Morphology. The latter is an analysis of the Kahnawake dialect.

[3] There are two "feminine" pronominal prefixes in Mohawk. One denotes female humans and indefinite persons (feminine indefinite) while the other denotes female humans, animals and inanimate objects (feminine neuter). The choice of prefix in reference to female humans follows certain general rules of social etiquette, contingent upon specific age, status and affectional relationships. Idiosyncratic usage also is a factor.

[4] The reflexive is realized in a number of allomorphic shapes: -atat-, -tat-, -atak-, -tak-, -atate-, -tate-, -atan- and -tan-. Similar allomorphy operates for the semi-reflexive: -at-, -t-, -ak-, -k-, -ate-, -te-, -an-, and -n-. Briefly, prefix-initial -a- is lost following vowels; an epenthetic -e- is added preceding consonants; -t → -k/_v; and -t → -n/__i. See Bonvillain 1972 and Beatty 1974 for additional details.

[5] In morpheme segmentation, the following abbreviations are used for verbal affixes:

Prefixes: rfx-reflexive, srfx-semi-reflexive, fact-factual, fut-future, opt-optative, trns-translocative, du-dualic, rep-repetitive.

Suffixes: asp-aspect, bnf-benefactive, cs-causative, prog-progressive, cont-continuative, rvrs-reversive, dst-distributive.

Also, nom-nominalizer, int-interrogative.

Finally, note that /ʌ/ and /u/ are nasalized vowels.

[6] Ambiguity between reflexive and reciprocal meanings is widespread in diverse languages. For example, Shibatani (1985:826) cites data from Spanish, Russian,

Quechua, Yavapai and Tetelcingo Nahuatl. Also, French se can have both senses when used with non-singular subjects. Finally, see Hollenbach for a discussion of ambiguity between "reciprocal and compositional readings" for the reciprocal construction in Copala Trique, a Mixtecan language of Mexico (1984:281-283).

[7] See Shibatani (1985:827-828) for examples of use of reflexives to denote spontaneous events or states in Spanish, French, Russian and Quechua.

[8] The Mohawk pattern of employing -at- for 'sell/lend' contrasts with some data from Japanese discussed by Watanabe (1984). In a paper on transitivity in Japanese, Watanabe noted that agents who give/sell "...have more control over events than the receiver, the buyer" (1984:246). Hence, in Japanese, constructions involving giving/selling employ nominative/accusative marking (typical of transitive clauses) on their two NPs while those involving receiving/buying take nominative/nominative marking (indicating lower transitivity). The distinction in Mohawk seems to stress a different semantic concept. Since -at- signals decreases in transitivity, its use in 'sell/lend' must be based on a different principle. Moreover, some verbs with -at- indicate subject's lack of control (eg. I had a misfortune, a misfortune happened to me). These facts suggest that the semantic basis for use of -at- in 'sell/lend' is more likely derived from concepts of movement as discussed in section 3.6.

Furthermore, Mohawk interest in movement related to prototypical place or direction is also attested in verbal prefixes called "cislocative" and "translocative." Among the complex contrastive uses of these prefixes is that the cislocative can signify movement toward prototypical place (i.e. ground level), eg. descend from upstairs, come up from a hole in the ground. In contrast, translooatives can denote movement away from ground level, eg. climb up into a tree, descend into a basement. See Bonvillain 1981 for further discussion.

[9] In a comparative note, the prefix ber- in Bahasa Indonesian (cited in Hopper and Thompson 1980:278) has quite similar functions to Mohawk -at-. It is used to detransitivize verb stems (to hang something up ⟶ to be hanging), to mark on-going states (to bear fruit), and to express reflexive and reciprocal meanings (to exchange rings, clasp hands, shave oneself).

[10] For discussion of phonological reconstructions for Proto-Iroquoian and subsequent sound changes in the

descendant languages, see Lounsbury 1978 and Mithun 1979.

[11] Note that the Northern morpheme -atat- is realized phonetically as -adat-/__C and -adad-/__V.

[12] In a study of Cavinena (a Tacanan language of Bolivia), Camp describes a reflexive construction which, in addition to reflexive and reciprocal meanings, also can be used without co-referential subjects and objects to express possessive, benefactive or locative roles (1985:55).

## REFERENCES

Barbeau, C.M. 1915. Classification of Iroquoian Radicals with subjective pronominal prefixes. Ottawa: Canada Dept. of Mines, Geological Survey, Memoir # 46.

Barber, E.J. 1975. "Voice--beyond the passive." Papers from the Berkeley Linguistics Society, 1:16-24.

Beatty, John. 1974. Mohawk Morphology. Occasional Publications in Anthropology, Linguistic Series #2. Greeley: University of Northern Colorado.

Bonvillain, Nancy. 1972. A Grammar of Akwesasne Mohawk. Mercury Series # 8. Ottawa: Nat'l. Museums of Canada.

_______________. 1981. "Locative semantics in Mohawk: time and space." Internat'l. Journal of Amer. Linguistics, 47:58-65.

Camp, Elizabeth. 1985. "Split ergativity in Cavinena." IJAL,51:38-58.

Chafe, Wallace. 1967. Seneca Morphology and Dictionary. Smithsonian Contributions to Anthropology, 4. Washington DC: Smithsonian Press.

_____________. 1970. A Semantically-Based Sketch of Onondaga. IJAL Memoir 25.

_____________. 1976. Iroquoian, Caddoan and Siouan. Current Trends in Linguistics, State of the Art Reports # 3. The Hague: Mouton.

_____________. 1979. "Caddoan" in Languages of Native America. (eds. L. Campbell and M. Mithun). Austin: University of Texas Press, pp. 213-235.

Cuoq, J.A. 1866. Etudes philologiques sur quelques

Langues sauvages de l'amerique. Montreal: Dawson Brothers.

__________. 1882. Lexique de la langue iroquoise. Montreal:J. Chapleau et fils.

Fagan, Sarah. 1988. "The English middle." Linguistic Inquiry,19: 181-203.

Hollenbach, Barbara. 1984. "Reflexives and reciprocals in Copala Trique." IJAL, 50: 272-291.

Hopper, Paul and Sandra Thompson. 1980. "Transitivity in grammar and discourse." Language, 56:251-299.

Jelinek, Eloise and Richard Demers. 1983. "The agent hierarchy and voice in some Coast Salish languages." IJAL, 49:167-185.

Keyser, Samuel and Thomas Roeper. 1984. "On the middle and ergative constructions in English." Linguistic Inquiry, 15: 301-416.

Lounsbury, Floyd. 1953. Oneida Verb Morphology. New Haven: Yale University Publications in Anthropology, Memoir # 48.

________________. 1978. "Iroquoian languages." in The Northeast, Handbook of North American Indians, vol. 15 (ed. B. Trigger). Washington DC: Smithsonian Institution Press, pp.334-343.

Mithun, Marianne. 1979. "Iroquoian." in Languages of Native America. (eds. L. Campbell and M. Mithun). Austin: University of Texas Press, pp. 133-207.

Reyburn, William. 1953. "Cherokee verb morphology I, II." IJAL 19: 172-180, 259-273.

________________. 1954. "Cherokee verb morphology III." IJAL, 20:44-64.

Rood, David. 1976. Wichita Grammar. New York: Garland Press.

Rudes, Blair. 1981. "A sketch of the Nottoway language from a historical-comparative perspective." IJAL, 47:27-49.

Shibatani, Masayoshi. 1985. "Passives and related constructions." Language, 61:821-848.

Thompson, Chad. 1989. "Pronouns and voice in Koyukon

Athapaskan." IJAL, 55:1-23.

Walker, Willard. 1975. "Cherokee." in Studies in Southeastern Indian Languages. (ed. J. Crawford). Athens: University of Georgia Press, pp. 189-236.

Watanabe, Yasuko. 1984. "Transitivity and evidentiality in Japanese." Studies in Language, 8: 235-251.

Williams, Marianne (Mithun). 1976. A Grammar of Tuscarora. New York: Garland Press.

Woodbury, Hanni. 1975. Noun Incorporation in Onondaga. New Haven: Yale University PhD. Dissertation.

# PROTO-ALGIC V: Doublets and their Implications

Paul Proulx

Abstract: Accumulating evidence suggests that a recalcitrant problem first addressed by Howard Berman - how to explain an apparent correspondence among PA *t, W t, and Y ? - is best solved by reconstructing Proto-Algic doublets. These doublets suggest dialect mixing before the breakup of Proto-Algic society, with frequent elements commonly manifesting the prestige-dialect innovations. The combination of *? with adjacent consonants, once e-dropping is taken into account, explains the glottalized consonants in several words - but others remain and should be transcribed.

## 0. Introduction

Ives Goddard, in describing Proto-Algic as being 'at the limits of our perception, where only dispersed fragments of the protolanguage can be perceived' (Goddard 1991:65), has wonderfully captured the essence of the challenge inherent in its reconstruction - and no aspect of that reconstruction justifies his observation better than the present problem. Work on such problems soon makes clear the nature of the scientific enterprise, teaching the researcher the need for patience and the constant reassessing and improvement of hypotheses. There is no quick fix, short cut, magic methodology, or the like that will do the job. Yet Goddard to some extent overstates the difficulty: gradually some of the fragments begin to fit together, honest debate sharpens insights, and the emerging picture comes ever into clearer focus.

The first paper in this series,[1] in which I presented an initial sketch of Proto-Algic phonology, was researched in the late 1970's and early 1980's - the same period during which Howard Berman was putting together his own sketch (Berman 1982a, 1984). Despite some correspondence between us, we evolved surprisingly different proposals.

This is a reflection of the fragmentary and often confusing and/or ambiguous nature of some of the data. However, it has the salutary effect of highlighting what Hockett (p.c.) calls the

*Kansas Working Papers in Linguistics*, Volume 19, Number 2, 1994, pp. 115-182

residue of reconstruction. At any stage of reconstruction, according to Hockett, there will be some more or less promising matchings which do not fit the researcher's current model. Like the discard pile in a game of cards, they lie there on the table waiting for someone to find the opportunity to put them to good use. The more residue available, the better the opportunity to make progress on the reconstruction. And highly divergent alternate hypotheses do generate copious residue for each other - for what is a cognate set under one proposal is only close (or not so close) to being one under the other. Moreover, this particular residue is of especially worthy note as it comes with the certification that at least one of one's colleagues considers the matchings to be true cognates.

In theory, of course, the same scholar could entertain a variety of alternate hypotheses (and probably does so early in the game). Due to the enormous complexity of language and of the reconstruction of older protolanguages, however, a scholar soon narrows down working hypotheses to keep from being overwhelmed by the enterprise. Alternate hypotheses become the work of other scholars. In this paper, I turn to some of the residue Berman and Goddard have generated for me and show how some of their insights can be harmonized with my own.

Despite this partial harmonization, however, there remain radical differences between my own proposal and Berman's - essentially accepted by Goddard (1991:64-65) - which make it hard for me to discuss his insights without to a great extent reinterpreting them in terms of my own understanding. In particular, Berman is persuaded of the Ritwan Hypothesis (see appendix B), while I am not. The reader wishing to experience Berman's point of view (or Goddard's, for that matter) is therefore referred to the original sources.[2]

Throughout this paper, I will be mentioning consonant and vowel grades, and archaic ablaut, which manifest as follows. First, Proto-Algic has four sets of sound symbolic equivalences among phonemes, where grade 1 is neutral, grade 2 has diminutive and/or meliorative connotations, and grade 3 (where it is found) augmentative and/or pejorative ones. When there is no grade 3, its meanings are reassigned to grade 1 (at least in the case of the first set). The grades in Proto-Algic, and their unconditioned reflexes, are:

| | GRADE 1=2=3 | PA | Wiyot | Yurok | |
|---|---|---|---|---|---|
| (1) | **t=c=ĉ | *t=s=ŝ | t=c=ĉ | t=c | (Proulx 1984:175) |
| (2) | **s=ŝ | *ł=ŝ | s=ŝ | s=c | (Proulx 1984:175) |
| (3) | **l=r | *l=ŝ | l=ŝ | r=s | (Proulx 1991:sec.3.1) |
| (4) | **e=a | *e=a | --- | e=ɨ | (Proulx 1984:175) |

That is, grade 1 **t matches grade 2 **c and grade 3 **ĉ, and so forth. Grade 2 of **s and **t coincide in Yurok, and backformations from this grade explain the handful of secondary matchings of **s and **t (the latter serving as the new augmentative since grade 3 has been eliminated in Yurok). Y sesomen- 'scratch' and tetomen- 'scratch repeatedly' is a good example.

Second, archaic ablaut consists of the replacement of **e by **a: in verb stems, dependent nouns with matching verb medials, and what look like nominalized verbs. Its function (where recoverable) is to imply iteration, habitual action, and the like. For examples, see Proulx (1984:sec.1.3, and nos. 016, 018, 041, 077, 078, 105).

Goddard (1991:fn.17) has claimed that reconstructions involving such grade variation are implausible. However, this ignores the well-documented grade variation in both Wiyot (t=c=ĉ, s=ŝ, l=r, see Teeter 1964:21-22) and Yurok (t=c, l=r, e=ɨ, see Haas 1970, Hamp 1970, Berman 1981:259, and Proulx 1982b) - as well as the somewhat less obvious evidence from Algonquian (*t=*ĉ, *s=*ŝ, see Goddard 1975:74, 79; and *e=*a, see Siebert 1967:8, 13, 48, and 1975:11, 51, 52, 59, 101, 114, 130, 170, 184).

New or significantly revised Proto-Algic reconstructions begin with number (359); numbers 001-358 are found in Proulx (1984, 1985a, 1991, 1992). A bare number after a Proto-Algic reconstruction or reference to one, or after an Algic cognate set, refers to this numeration. The cognate sets underlying Proto-Algonquian reconstructions for which no source is given will be found with the earlier Proto-Algic reconstruction whose number is cited.

## 1. The history of the problem

Early in my research, I came upon a potential cognate set PA *ŝeki-, W tikɨl-, and Y ?ahk- 'urinate' 099. I passed it on to Howard Berman, who after initial skepticism went on to find more examples of Y ? apparently corresponding to W t and PA *t (or one of their grade variants), as well as to Y t in doublets: PA *t(a:)- 'exist', W t(a)- 'stay, dwell, remain', Y ?(o)- 'be, exist, be born' and the near doublet Y to:?m- 'be together in a group' (with -o:?m 'incremental plural'); PA *ahŝam- 'feed', W t- 'feed, give food to', Y ?- 'give'; PA *net-, W dut-, Y ?ne- 'first person' (and similarly with the second and third); PA *nyi:ŝ(wi), W dit- 'two', Y ni?(iyeł) 'two (human beings)' (Berman 1982a:417ff). He also compares the final segment in W wɨtkɨdát, Y ?wɨłkɨ:? 'bone'; and W ta 'durative and articular preverb', Y ?o 'locative preverbal particle'. He suggests a second doublet, Y -e?wey 'face' beside tewey 'forehead', where the initial Y e in -e?wey is a link vowel (as shown by the allomorph -o?wey, see Proulx 1985b), while that in tewey 'forehead' would presumably have been dropped in -V?wey (Berman 1982a:fn.12).

Meanwhile, I too found some further examples of the same sort: W dɨ́tɨbák 'for two days', Y na?amo?- 'be somewhere for two days' 065; and PA *no:ĉpen-, Y no?p?en- 'pursue' 098. I explained the former morphologically, the latter and 099 by postulating clusters and divergent simplification. I explained the prefixes as follows. Proto-Algonquian has two sets of prefixes used before nouns beginning in vowels, e.g., *w- and *wet- 'third person'. Blackfoot has eliminated the first set, everywhere substituting the second (e.g., ot-). Proto-Algic has two comparable sets (Proulx 1984:169, 198), but Yurok has only one. I assumed that Yurok, like Blackfoot, had analogically leveled - but in the opposite direction (replacing the second set by the first, e.g., Y ?w- beside PA *wet- would respectively reflect **w?- and **w?et-). I further assumed that the leveling out of the **et extended to some cases where etymologically the **t had belonged to the stem: PA *wehtehkwani 'branch', Y ?weskwen 'small branches, twigs' 038; PA *-atayi, Y -eyah 'belly' 095; and PA *watapya 'fine root used in sewing', Y ?wohpeg 'spruce root' 102.

At the time, these explanations seemed more plausible to me than Berman's solution, which was to equate Y ? with W t and PA *t (and their grade variants) in items of this sort and assign both this correspondence and the regular *t+t+t one to the same proto phoneme *t - for neither of us could find a phonological conditioning environment which would put these two correspondences in complementary distribution. And indeed, no such environment has ever been found. Consider the following contrasts of **T versus **t (and **K versus **k), where the apparent *t+t+? and *k+k+? correspondences (including grade variants of t) are respectively written **T and **K:

| | |
|---|---|
| **niTema:K- 'two days' 065 | cf. **pitekwli 'basket' 111 |
| **no:Tpen- 'pursue' 098 | cf. **ketp- 'thick (cloth)' 040 |
| **Tik- 'urinate' 099 | cf. **ti?ema:? 'elderberry' 066 |

Berman's solution forces one to postulate dialect mixing of some sort - and dialect mixing can be a too-easy solution to every apparent deviation from regular phonological correspondences. Therefore, one properly arrives at it by differential diagnosis, when all other plausible explanations have been tried and found wanting. By this test, his conclusion seemed premature.

Moreover, the prefixes are important here, and Berman's suggestion that, for example, **net- gave Y ?n- 'first person' by regular sound change seemed unconvincing to me. His explanation was that Yurok glottal catch (including that from *t) could move backward in a word, as in Y no?nowos 'fetch!' (stem nonow- 'to fetch' plus imperative singular -?os). However, it doesn't move backwards in Y no?p?en- 'pursue' beside PA *no:ĉpen- 098, so it appeared that a different phenomena was involved than in the imperative. I also resisted the idea of a *t > ? sound shift before vowels as somewhat unnatural (though not impossible).

Nevertheless, Berman (1984:fn.11) found another Yurok doublet (Y ?ahtemar and dependent -?a? 'drawing, book, newspaper'). Later he proposed two more cognate sets: W dat- 'be large', Y no?(op)- 'be tall, high (round things, mountains, trees)' and no?(omek)- 'be long (worms, ropes)' (Berman 1990:432); and Y ?yekw(ł) and PA *wetkw(e:wa) 'maggot' S248 (citing the latter as PA *o:tkwe:wa. Goddard (1991:65) proposed PA *-t, Y -? 'third person subject'. Finally, I found one more: PA *wełkiĉyi 'pipestem', Y ska? 'pipe scabbard' 331 (Proulx 1992:34). It was time for me to take another hard look at the question, but first I had to recognize its relationship to two others.

## 2. Some losses of **e

In my 1984 sketch, I reported some losses of **e 'between an obstruent and a consonant other than **g' in a first syllable in Wiyot and Yurok (p.186). I also reported loss of **e between **n and a stop or affricate in PA and Yurok (p.187), though only word medially, and of initial **ne before a single stop in Yurok (p.197). Finally, I pointed out that the environment for **e-loss had to have been broader (at least in Wiyot) to account for the vowel-less prefixes (p.198) - though the details are largely obscured by analogical restorations.

Additional evidence now leads me to also include the **e of the Yurok prefixes among the losses. For example, **e (and preceding **w) are lost in Y ska? 'pipe scabbard' beside PA *weɫkic̑yi 'pipestem' 331, and in Y lo?og beside W wɨràg 'coals'. Other examples of the loss of initial **we are: 074, 102, 279, 289, and perhaps 280 (if its first **a had an e-grade).[3]

The loss of **e in a first syllable makes Berman's proposal for Y ?n- from **net- much more attractive: presumably, **net > *ne? > *n? > Y ?n or something of the sort. No backward movement across vowels needs be assumed.

## 3. The Reflexes of **?, and Some Proto-Algic Doublets

The distribution of ? (subtracting known secondary sources) in the three branches of Algic is as follows. In Wiyot, it's found before a continuant and in word final position. In Yurok, it's found everywhere except before a consonant (ignoring ? as a laryngeal increment and preglottalization of sonorants and erstwhile sonorants, e.g. ?r < Pre-Y *r?). In PA, most cases of *? are probably secondary - but evidence of this is often lacking. It is found as the first member of a consonant cluster and, written *h, between a preceding vowel and a following vowel or semivowel (with a morpheme boundary intervening in the known cases).

**? is never preserved before a stop in any of the languages. PA and Wiyot evidently preserve (or innovate) it before **w: PA *-a?w, W -u?w and -utw (Teeter 1964:72-73) 'by tool, instrument, medium' 187, 369. If PA *-h 'abstract transitive final' (Bloomfield 1946:sec.86) is in origin a doublet with specialized *-t of transitive inanimate verbs (**-Vt 'nonpersonal object' 178), this provides a good example of it occurring before a morpheme boundary plus, generally, a vowel.

**? is preserved (or innovated) in word final position in Wiyot and Yurok, and perhaps between unlike vowels in a word final sequence in PA: PA *-t, W -i?, Y -o? 'third person subject' 300; and perhaps (with a different inflectional vowel) PA *-ahi 'locative 1' (Proulx 1988:322) - possibly related to Y ?o 'locative preverbal particle'.[4]

**? drops in Wiyot and Algonquian in absolute word initial position and between vowels within the root. Word initial examples are W kɨ́?y (ukiyɨbɨ?l 'her hat'), Y ?ekah (**?eki?yemi 'hat' 047); and PA *e:nsa, Y ?ɨncɨh 'abalone shell' (**?e:neca 'bivalve shell' 084). An intervocalic example is **wenli?a:gwi 'coals, charcoal' 289: Y lo?og, W wɨràg, cf. PA *-anŝye:wi.

Of course, these are only rough preliminary statements of the environments. PA *aɬemwa 'dog' A205 beside deverbal *-a?ɬemw A176 of the same meaning suggests that, in Algonquian, position in the word may be as important as the phonological environment. Or, the glottal catches may be secondary in such cases.

However, some of the examples just cited strongly suggest that Berman's Yurok doublets are of Proto-Algic antiquity. Consider PA *-a?w, W -u?w and -utw (Teeter 1964:72-73) 'by tool, instrument, medium'; PA *-t, W -i?, Y -o? 'third person subject'; and PA *tahkon- (a-grade) 'grasp (it)', W khɨ́n- 'grasp, hold', Y tekon- and ?ekon- 'hold onto something'. These items show presumed reflexes of **? in Wiyot and PA corresponding to those of **t elsewhere (or in doublets in the same language) - something impossible if the dialect mixing had taken place only within Yurok. Consider also the doublets PA *ta:n- and *a:n- 'WH-' (Proulx 1991:152), and Cree ce:skwa: (Cree diminutive grade c from PA *t) and e:skwa: 'still', as well as nama ce:skwa: beside name:skwa: 'not yet'. Compare also W -u?w and PA *-etwi 'reciprocal' (Bloomfield 1946:sec.72, with the intransitive final *-i).

All of the foregoing is surely sufficient evidence to warrant serious consideration of Berman's hypothesis, modified to reflect the greater antiquity of the doublets.

## 4. Testing the hypothesis

So far we have seen evidence for several Proto-Algic doublets. Those longer than a single segment are (where for present convenience **T is written for a **t + **? pair, and **C and **Cˇ respectively for its grade 2 and 3 variants):

(359) *Tik- 'urinate' 008, 099: PA *ŝeki- (grade 3, with mediopassive *-i) A1980, W tikɨl- (with stative -ɨl), and Y ?ahk-.

(360) **Tahkon-, **Tehkon- 'grasp, hold onto something with the hand' 064, 019: (a) PA *tahkon- 'grasp', (b) Y tekon- 'hold onto something'; Y ?ekon- 'hold onto', W khɨn- 'grasp, hold'. See sec.6 for the origin of the aspiration in Wiyot.

(361) **weskiTyi 'hollow tube used in smoking' 331: PA *wełkiĉyi 'pipestem' (C oskiĉiy, O okkiĉ), Y ska? 'pipe scabbard'. PA has regular mutation of t to ĉ before y.

(362) **wehTelkweni, **wehTelkwani, **wehTerkwani 'branch' 038: (a) PA *wehtehkweni, (b) Y ?wełkun 'heavy limbs', PA *wehtehkwani, (c) Y ?weskwen 'small branches, twigs'.

(363) **-eThey, **-aThay 'belly' 095, 145: medial PA *-eĉy, medial W -ith, dependent Y -eyah (with **-i 'inanimate', see Proulx 1991:133), (b) dependent and medial PA *-atay. A fully inflected form, showing movement of Pre-Yurok *? back onto the prefix, is **neTheyi, **naThayi 'my belly': Y ?neyah, PA *natayi.

(364) **weTaphega and **weTa:phega 'root, fine ... used in sewing (usually spruce)' 102: PA *watapya S170, W tàp 'spruce root' (stem táph-), Y ?wohpeg 'spruce root'.

(365) **n-, **net-, **ne?- 'person, first...' 128: (a) PA *n-, W d- or zero, Y ?n- (with ? analogical to the second by-form), (b) PA *net-, W dut- (vowel analogical to third person prefix), Y ?n-, (c) PA *ne-, W du?- (vowel analogical), Y ?ne- (? analogical).

(366) **K-, **Ket-, **Ke?- 'person, second...' 129: (a) PA *k-, W H- (the aspiration of an immediately following obstruent) in kin nouns specifying most senior consanguines, and W kh- otherwise (aspiration secondary), Y k?- (with ? analogical to the second by-form), (b) PA *ket-, W khut- (vowel analogical to third person prefix, aspiration secondary), Y k?-, (c) PA *ke-, W khu?- (vowel analogical, aspiration secondary), Y k?e- (? analogical).

(367) **w-, **wet-, **we?- 'person, third...' 130: (a) PA *w-, W w-, Y ?w- (with ? analogical to the second by-form), (b) PA *wet-, W hut-, Y ?w-, (c) PA *we-, W hu?-, Y w?e- (? analogical).

The short form (set a) is used in dependent nouns. The glottal variants of the long forms of the personal prefixes (set c) had evidently been adopted throughout Proto-Algic before independent elements which began in consonants - leaving their stop counterparts (set b) only for the small minority of cases where the following independent element began in a vowel.

The history of Wiyot second person prefixes begins with a specialization of **K-. The glottal variant is used with most senior consanguines, the stop variant elsewhere. Next, after the loss of the connective **e, the glottal catch metathesizes with a following obstruent and turns into aspiration. Finally, that aspiration spreads to the stop variants of the prefix, creating a blend. In Yurok, the glottal (which is found in allomorphs of all three prefixes) is analogically extended from set b to sets a and c. Whatever its source, then, glottalization (and resulting aspiration) have tended to spread among the variants of the personal prefixes.

The final glottal catch in set c is only preserved (in careful speech) before sonorants in Wiyot. Despite Teeter's ambiguous statement of the matter (Teeter 1964:79), his own chart showing the distribution of consonants (p.16) as well as his examples (pp.79-80) shows they disappear without trace before obstruents. Perhaps they originally metathesized with the following obstruent and produced aspiration - but if so the resulting irregularity in the stem, lacking a morphological function, was leveled out.

(368a) **-eTew 'reciprocal': PA *-etwi (Bloomfield 1946:sec.72, with the intransitive final *-i), W -u?w, and Y -ew. Wiyot shows sporadic rounding (see Proulx 1984:181). Y ? has presumably migrated onto a previous element and been lost, after the loss of the initial **e (**e? ---> *?).

(368b) **-eCew 'reflexive' 192: PA *-eswi A333; Y -ew (in noninflecting verbs) R48-49. Compare W -w, -iw T74 (lacking glottal catch). Several of the proto subject endings, which would follow this element, have **K or **T as their consonants. Had their glottal variants followed Pre-W *-i?w, its glottal catch would have dropped by laryngeal elision (Teeter 1964:26) - if this type of elision dates back that far.

(369) **-VTw or **-VTew 'by tool, instrument, medium' 187: PA *-a?w, W -u?w and -utw (Teeter 1964:72-73). PA also has *-ah 'by tool, instrument, or medium' A76 in TI verbs and their derivatives, by analogy with the many TI stems which lack the *w or *aw of their corresponding TA stems.

(370) **ehT- **ahT-, **ahCam-, **ahCˇam- 'give (especially food)' 011: (a) W t(u)- 'feed' (see Berman 1982a:fn13), (b) Y ?(o)-, (c) W łcłb- 'give to' (grade 2), and (d) PA *ahŝam- 'feed' (grade 3, with final *-am, see Bloomfield 1946:sec.82). The last two proto stems have stem root extension **-Vm 355.

(371) *no:Tpen- 'pursue' 098: PA *no:ĉpen-, Y no?p?en- 'pursue'.

(372) **niTeThey-, **niTeCˇhey- 'two humans' 036: W dłt-

'two' (grade 1), Y ni?iy-; PA *ni:ŝwi 'two' A1630 (with root extension **-Vw 357). The root is **niT-, followed by grades 1 and 3 of medial **-eThey 'belly' 095, 145, 363. **y and **ey evidently drop in Algonquian after PA *ŝ from **ĉ and **ŝ (cf. 086). Evidently too, Yurok has haplology - or perhaps **TeT ---> *?e? ---> *?? and thence Y ?. The loss of aspiration in Wiyot is evidently analogical from word-finala position where it is regular (e.g., in phrases like dit ba hɨ́lɨd '20').

(373) **niTema:K- 'two days' 065: W dɨ́tɨbák 'for two days', Y na?amo?- 'be somewhere two days'.

One of the main tests of a rightful hypothesis is productivity, that is, its ability to generate new insights. If Proto-Algic really had doublets of the sort envisaged above, knowledge of this should allow new or improved reconstructions - and perhaps new insights into the relationships among these and/or existing ones. Those new reconstructions longer than one segment follow:

(374) **na:T- 'large' (Berman 1990:432): W dat- 'be large', Y no?omek- 'be long (ropes, etc.)' (with **-Vm 'root extension' 355, and **-Vkh 'long thing' 155), no?op- 'be tall (round things)'. The reduplicated form Y no:no?op- does not show a "laryngeal increment" (pace Berman, and cf. fn.2 below). It presumably reflects Pre-Y *no?ono?op- with contraction of *o?o to Y o:. Compare Y mɨkwɨɬ 'peak' and reduplicated mɨkwɨmɨkwɨɬ 'peaks' R14 for this type of reduplication, and Y co?one?n, and co:ne?n 'four (body parts, etc.)' R88 for sporadic contraction of o?o to o:.

(375) **weTewi, **wegeTawi 'flesh, body, (her)...': (a) Y tewon 'flesh' (with nominalizer **-Vn 296), ?wetew 'her flesh'; (b) PA *wi:yawi 'her body' A2265, and diminutive *wi:yawehsi 'meat, flesh' A2264. This noun is doubtless the source for the reciprocal and reflexive suffix **-eTew, via the meaning 'self' (see Bybee, Pagliuca, and Perkins 1990:36). The meaning 'self' turns up in M ne:yaw 'my body, person, self', cognate K niiyai and nɨ́ai 'myself', and derivationally *-iyaw in C tipiyaw 'in person' (f+w C weyu tipeyuwê 'herself', with ^ for macron).

(376) **peTk-, **peCk- 'catch fire': (a) Y pek?onoc- 'set on fire' (with Y -Vnoc 'causitive', as in roykenoc- 'render fat, force out liquid' beside royk- 'flow, drip, trickle', see Proulx 1985b:123), (b) PA *pesk- A1855. See sec.6 for the position of Y ?.

(377) **Ta:wal- 'be around': Y ?o:le?m 'be, exist' and Y ?o:lo?op? 'she stands' R271, W tá?lib 'it lies there' (with **-Vm 'third-person subject' 295. Compare **a:wal- '(go) around'

respectively in Y ho:le?m 'go travel, be around, fare', Y ho:ro?op? and ho:ro?opep? 'she runs around' (Proulx 1985b:130), and W hálit 'you go, walk'.

(378) **weTehpehKa, **weCehpehKa '(her) backbone': Y ?up?ɨh, PA *wesehpehkwa (M osɛ:hpɛhkon 'her backbone', with noun final *-w [see Bloomfield 1946:sec.57, 61]). **weTehpehKa gives Pre-Y **w?ep?a, and thence, with vowel harmony, **?wap?a and Y ?up?ɨh.

The general Algic pattern is for the part-of-whole noun to be homophonous with its corresponding dependent noun inflected for the third person (as in this case),[5] e.g., PA *wetkanĉege 'hoof, claw; her fingernail' 39, 134. For other examples, see Proulx (1992:36) and also 74 and 112. Often, as in 'backbone' above, attestation in the daughter languages is incomplete (see 035, 090, 091). In some cases (022, 322, and, for Wiyot, 090), the part-of-whole noun is treated as an ordinary independent noun, with allocative prefixes simply added.

The root **tekw- 'together, grouped (i.e., at the same place)' 380 may be related to the particle **?ékw 'just exactly at the same time or place, or in the same manner' 379. The semantic specialization in the latter pair (if such it be) was already present in the proto language, suggesting that doublets were not new.

(379) **?ekw- 'same, just exactly the time, manner, thing': PA *aŝikwa (NiO azhigwa 'at this or that time', with *aŝi- 'like that' 255 prefixed after loss of the initial syllable), W kwi- preverb 'on arrival' T86, e.g., kwi-kłám 'I see her when I get there', kwi-khɨnápił 'she just bites', Y ?ekw 'that's just exactly what or how' R140. Compare PA *e:kwe:la:ki 'exactly then' (pC e:kwe:ya:k, nC e:kwe:na:k); PA *e:kwa: and *me:kwa: 'while, in the act, at the place, at the time': NiO me:gwaa 'while, at the same time', pC me:kwa: 'while, in the act or place', wC âkwu 'now, whiles, whilst' (Faries), M mɛ:k- preverb 'engaged in, while the action goes on'. If related, these additional Algonquian forms have initial change and obscure affixes.

(380) **tekw-, **takw- 'together, grouped': PA *takw- 'joined, along with, together, grouped' A2016-2017, W ĉakw- in ki-piwɨ-ĉakwɨlɬwɨ?n 'one doesn't gather them much' (grade 3, with archaic ablaut), Y tek- 'together', e.g., teko?opi? 'it's burned on' (-o?op 'burn'), teku:ne?m 'they grow together' (-u:n 'grow'). Compare W takwun or taku 'they' T100, and W tɨla-ĉkwłayúwi? 'how they hook eels' T47.1.

Finally, another set of apparent doublets in **T turns out to have a different origin. When a dependent noun stem beginning in **t has a counterpart used as a verb medial, the latter generally lacks the initial **t.[6] In these cases, the **t is a structurally-required prefix receptor (Proulx 1992:19), normally absent after a verb root (though extended there analogically in some cases in the daughter languages, see Proulx 1985a:66-68).

## 5. Some **? vs **k Doublets

As we have seen, two of the **T doublets are also doublets in **K: **niTema:K- 'two days' 65, 373 (W dɨ́tɨbák 'for two days', Y na?amo?- 'be somewhere two days') and **weTehpehKa 'backbone' 378 (Y ?up?ɨh, PA *wesehpehkwa 'her backbone'). A search has turned up several more. Those of more than one segment are as follows.

(381) **-VKa:p 'stand': PA *-ka:pawi A516 (with extension **-Vw 357), Y -o?op 'stand, run'. Examples are PA *kwe:łkika:pawiwa 'she turns as she stands' A1115, and Y ?o:lo?op? 'she stands' R271 (with ?o:l- 'be around', see 377).

(382) **Kiyew- 'in a circle, around, about': PA *ki:w- 'around, turning, returning' A993-1017, W hiyɨw- 'around, in a circle'. Examples are: W hiyɨwéliliski?l 'circle around the sun', kawi?yɨwéta?l 'they go around and around' (kaw- 'begin', -Vtal 'go'), ta-yɨwa?nɨ́talił 'she's walking around in a circle' (ta- 'progressive', **-Vn 'root extension' 356), hiyɨwa?nākw 'I'm dizzy'; PA *ki:we:wa 'she goes home, back' A1006, *ki:wi?ta:pahta:wa 'she runs in a circle, runs around something' A1013, *ki:waŝkwe:- 'be dizzy, silly' A1000-1001.

(383) **no:K- 'such a distance': W nuk-, Y no?- in no?om- (incorporating postradical *-Vm 355), and no:ł 'far, long time' (incorporating locative *-Vł). Examples: Y kus no?omekik? 'how far did this water flow from?', Y kus no?omenek?w 'how long was it submerged?', Y kus weno?omekik? 'where is the water up to (e.g., in an aquaduct)', Y kus no:ł 'how far?, how long?' R.232 (with contraction of o?o), W ta-nukɨ-bɨ́wi?yɨwił 'they eat it for a long time' 54-f, W łé-dukikwłił 'she cried a long time' 14-g, W ta-núkɨk 'after a long time' 47-g.

(384) **-Vŝ-eKoK, **-Vs-aKoK 'I-you pl.': Y -Vc-?o? R70 (e.g., nekcenic?o? 'I meet you pl.' R71), PA *-eł-akok (Proulx 1990:111).

(385a) **neKila(wa) 'first person singular personal pronoun': **nekila(wa) gives Y nek (shortened, **e retained in monosyllable); **ne?ila(wa) gives PA *ni:la(wa) A1610-1611 (where presumably **e?i gives PA *i:), and W yłl (presumably via **neyil).

(385b) **KeKila(wa) 'second person singular personal pronoun': **?ekila(wa) gives W khìl (where the initial **?e produces the aspiration after e-dropping as in possessed nouns), and **kekila(wa) gives Y kel- (presumably by haplology, or because *kk simplifies to k after e-dropping). **ke?ila(wa) gives PA *ki:la(wa) A896. This reconstruction (385b) replaces 260. Evidently, Algonquian gets its prefix **k(e)- 'second person' from a k-dialect of Proto-Algic, but its personal pronoun from the ?-dialect (the reverse of Wiyot). A similar union of mismatched elements is seen in **kega:rKa 'gull' 135 (PA *keya:ŝkwa with noun final *-w, Y kego?s-neg with neg 'animal; one who always eats'). Evidently, the final **-Ka is a nominalizing suffix in origin, rather than part of the root, and its glottal varianat is attested in Yurok while its stop one is found in Algonquian.

(385c) **weKila 'third person singular personal pronoun': **wekila gives W kwilɨ́?l (with obviative -ɨ?l), and Y kwelas (with obviative -Vs); **we?ila gives PA *wi:la A2233. This reconstruction replaces 262.

Items 385b-c suggest that the reflex of **i before a sonorant in Yurok is e. Another example is **thigw- 'out, through an opening' 288 (W thig-, Y tew[oy]-, PA *taw-). Yurok has the root extension **-Vy 358, and PA has grade 2-3 *a from its regular reflex *e of **i in a first syllable. This is plausible, for before an obstruent **i gives Y a (Proulx 1984:181), and Y a and e are barely contrastive (see Berman 1982a:413). Since Y a and e are also the (partly overlapping) reflexes of **e (see Proulx 1992:14-15), it seems likely that **i and **e merged in Yurok - and that the apparently slightly different environments for Y a and e in the two cases are due to the vagaries of attestation. The partial overlap in the distribution of Y a and e may be due to differences of stress as suggested by Berman (1982a:413), or perhaps to dialect mixing.

If **i and **e merged in Yurok, as I believe they did, then the following items and their variants (all with Y i) should be reconstructed with long **i: rather than short: **leski:m- 'scold' 41, **i:m- 'hurriedly' 123, **i:?meli 'below' 124, **-i:s 'paddle' 180.

Improved reconstructions in **K:

(386) **e?m?i:Ka 'pigeon' 121: W ɨ?mɨ́k, Y e?mi?.

(387) **-so:Kani 'hip' 133: PA *-ɬo:kani, and Y -to? [with shortening, and grade 1 t for s].

(388) **naKh- 'be skillful' 229: Y nekomur- 'swim well' (e-grade, with **-Vm 'root extension' 355), W dɨkhw- 'be good at'

(with **-_Vw_ 'root extension' 357), PA *_nak_- and *_nek_- (with various root extensions); but also PA *_nah_- (broad phonetic *-[_na?_]) in C _naha:piw_ 'she sees well' (with medial *-_a:p_ 'look' A45) and *_neh_- in M _n‹he:htaw_ 'she makes it skillfully' beside _ese:htaw_ 'she makes it so'.

The discovery of **_K_ also explains the alternate plurals of _we?yon_ 'girl': _we?yono?_ and, with the diminutive grade of vowel, _wɨ?yɨnɨk_ 'girls' R23. It is even possible that PA *-_Vki_ 'animate plural' and *-_Vhi_ (broad phonetic *-[_V?i_]) 'obviative animate plural' are by-forms of a plural suffix **-_VK_.

## 6. _?-movement, ?-deletion_

Glottalization of a Yurok consonantism is according to Robin's (1958:38) rule: _?_ is found on all stops, and between two continuants (e.g., _no:yc?k?w_ 'she eats as a guest', stem _no:yckw_- plus -_?_ 'third person singular'). With one exception, **_?_ from **_T_ always glottalizes an immediately adjacent consonantism in Yurok when there is one (due to the loss of **_e_). Whether that consonantism precedes or follows evidently makes no difference. Most examples are of the former type (synchronically, Yurok is a suffixing language; diachronically, e-loss is most common in a first syllable in Yurok, frequently before **_T_). Diachronic examples are:

**_wehTelkwani_, **_wehTerkwani_ 'branch' 038, 362: Y _?wełkun_ 'heavy limbs', Y _?weskwen_ 'small branches, twigs'.

**_neTheyi_ 'my belly' 095, 145, 363: Y _?neyah_.

**_weTa:phega_ 'root' 102, 364: Y _?wohpeg_ 'spruce root'.

**_neT_- 'person, first...' 128, 365b: Y _?n_-.

**_KeT_- 'person, second...' 129, 366b: Y _k?_-.

**_weT_- 'person, third...' 130, 367b: Y _?w_-.

**_weTehpehKa_ '(her) backbone' 378: Y _?up?ɨh_.

There is also at least one Wiyot example of ?-movement to the left, which produce aspiration of a preceding obstruent:

**_apiT_- 'tie' 162, 347, 389, 410-412: W **_ɨph_-.

I have found two cases of ?-movement to the right (with metathesis) in Yurok, and two in Wiyot. Note that the first _?_ in Y no?p?en- is the regular laryngeal increment, not the reflex of **T.

*no:Tpen- 'pursue' 371: Y no?p?en-.

**peTk- 'catch fire' 376: Y pek?onoc- 'set on fire'.

**Tehkon- 'grasp, hold' 360: W khɨ́n-.

**Keĉo:la 'thy maternal aunt' 048: W ĉhʉ̀l T81.

Ambiguously to the right or left:

**KeKila(wa) 'thou' (385): W khɨ̀l.

In the case of **Tehkon- 'grasp, hold', the reflex of the absolute initial glottal catch was presumably restored in Wiyot by analogy with those cases where it came before preverbs, as in kwi khɨnáphiɬ 'she bites on arrival' T86, 25-h (**?ekwi-?ehkon-a:p-eT-iɬ 'at-exactly-the-same-time, grasp-by-tooth-it-she'). In the case of **Keĉo:la 'thy maternal aunt', the morphological function of the erstwhile glottal catch has evidently preserved it.

As all Proto-Algic clusters of occlusive plus **k are simplified to k before Y w, *? from **T never glottalizes a following Y kw. In the following examples, I write the missing segment **T in the absence of evidence for **t (as in 390, where the initial **t guarantees the second). In 395, **T is guaranteed by the Wiyot cognate.

(389) **apiTeTk(w)-, **epiTeTk(w)- 'tie into a round shape, i.e., a bunch or bundle' 346: (a) Y pɨkwɨy 'hill, back part of house', W hɨphɨtk- 'tie roundish objects', (b) PA *petkw-, medial *-a:peŝk 'lump, sphere, knot' (Proulx 1992:54). Yurok has no initial p? in the lexicon, so presumably **piT ---> **pe? ---> **p? ---> p. The dependent stem would be analogical. PA evidently has haplology after **i gives *e in a first syllable. Possibly, Y Vy is from an unreconstructible nominalizer **-Vy, as suggested by **-Vy 'third person subject' and the general identity of Proto-Algic third-person and nominalizing suffixes (see Proulx 1992:25). Among the Micmac, at least, 'the back part of the camp...is up' (Rand 1971 [1894]:xxxii).

(390) **tetkw-, **tatkw-, **taĉkw- 'short, be ...' 079: (a) Y tkw-, (b) PA *tatkw- (F ĉahkwi 'short', O tɨkk(o:ya:) 'it's short', Ch e-tse?keso 'it's short', C tahkosiw 'she's short', A ĉee?ehixt 'she's short') and (c) *taŝkw- (M taski:w 'she's short').

(391) **leweTk- 'soft' 085: PA *lo:tk- S126, Y rewkw-. Cf. Y sewkwelum- 'be bruised' (from **rewetk- plus ?).

(392) **s?yoyeTkw-, **s?yoyeCkw-, **s?yoyeCˆkw- 'slip, be smooth' 086: (a) PA *ŝo:tkw- 'slip', Y syoykweɫ 'a slippery place' (with -eɫ 'place', Proulx 1985b:133), Y t?yoykwo?n 'it's slick' (grade 1 backformation, with **-Vn 'intransitive nondeliberate action' 177), (b) PA *ŝo:skw- 'slip'; (c) PA *ŝo:ŝkw- 'be smooth'.

(393) **kweyeCˆkweyehr-, **kweyeTkweyehl- 'whistle' 088: PA *kwi:ŝkwihŝiwa 'she whistles' A1121 (with the AI final *-i A453), Y kweykweyur.

(394) **yeTkw-, **yaTkw-, **ya:Tkw- 'curve': Y yekwohsok? 'I bend or fold it' NtBk 4:57, M wack- 'in a curved course, roundabout, by a detour' (reflecting **yatkw-, with Algonquian grade 3 consonant symbolism), M wa:k- 'bend, curve' (with archaic ablaut). The shift of **y to PA *w in word initial position is regular (see 'maggot' 318). The Menominee roots show the dissimilation of labials found in other words (Proulx 1984:178). Menominee examples are: wa:kepa:htaw 'she runs in a curved path' beside wackɛ:pahtaw 'she runs a curved course, by a detour', and wa:kesam 'she cuts it curved' beside wackɛ:sam 'she cuts it in a curve, cuts around it'. Yurok -ohs is a transitive final.

(395) **aleTewk-, **a:laTewk-, **a:leTewk-, 'be a cloud or shadow': (a) Y lewkw(enoɫ) 'be cloudy' (and perhaps Y rewkw[oɫ]- 'be misty'), PA *aletkw- (in Pe àlohk 'evening or night cloud', Mi alihkwatwi 'cloud', see S51), (b) PA *a:latkwatwi 'it's a cloud, it's cloudy' S51 (with II final *-at), (c) W halúkŝ 'shadow' (with nominalizer **-Vr). Compare Mc aluk 'cloud', which also rounds the **a. Presumably, **eTewk ---> *e?wk ---> *u?k ---> W ukh and then regular loss of sapirtion before a consonant. Similarly, **eTewk ---> *e?wk ---> *e?wkw ---> Y ewkw. The PA words contain medial *-atkw, *-etkw, deverbal from *watkwiwi 'cloud' S51 (with nominalizer **-Vw 304). For the archaic ablaut, compare 315 'projectile point'.

There are a few irregularities where one suspects a morphological origin, as in Berman's example of Y no?nowos 'fetch!' (stem nonow- 'to fetch' plus imperative singular -?os). Specifically, Y -ow 'do, act, be' (Proulx 1985b:130, 124) was probably added to many roots after the habit of glottalizing the previous consonantism had been established. Compare for example Y hu:nowoɫ 'they grow' (-oɫ 'they') and incremental (collective) plural hu:ne?m 'things grow' (-V?m 'incremental plural'), showing the root hu:n- 'grow' with and without -ow. Similarly, -ow is regularly lacking before the third person suffix -? (see the examples listed by Robins [1958:38]). Late addition of -ow is a plausible explanation for the glottalization of the previous consonant in imperatives in most cases, and analogy may be responsible for the cases like 'fetch' where the secondary nature of stem final ow is not clear.

## 7. On dialect Mixing

As we have seen, two of the **T doublets are also doublets in **K. In both cases, one variant has glottal reflexes of both while the other variant has the stops in both cases. Indeed, scrutiny of the reconstructions shows that reflexes of **T + **K in a single element must be: (a) both glottals, (b) both stops, or (c) respectively **? and **k.

This situation is somewhat reminiscent of that found in Cheyenne, an Algonquian language of the Great Plains. In Cheyenne, PA *k and *p have stop variants (reinforced with a prefixed laryngeal increment Ch h) and others it is convenient to call glottal variants (Ch ? - reinforced with a prefixed laryngeal increment consisting of a chameleon vowel when late in the word, zero elsewhere). For details, see Goddard (1978, 1988). Besides the three combinations of Proto-Algic (a-c above), Cheyenne has a fourth. The reflexes of *k + *p in a Cheyenne word must be: (a) both glottals, (b) both stops, (c) respectively ? and p, or (d) respectively k and ?.

Goddard (1978:75ff, 1988:335) has repeatedly made the puzzling claim that such a situation cannot result from dialect mixing. However, nothing could be a more normal, expected, and even typical result of it. Consider for a moment four groups of people speaking the same language, with at least two of them having high prestige in the society, another being under the influence of these two, and the last being relatively isolated. One of the prestigious groups modifies its pronunciation of k, the other that of p. The influenced group immitates both shifts, the isolated one neither. Later, under the impact of war, forced migration, or the like, their populations (and vocabularies) become mixed (and the erstwhile prestige factors largely lost). The resulting situation would look exactly like the Cheyenne one.[7]

The Proto-Algic case is only slightly different, in that only one innovating center is required - but with two innovations: one (the modification of **t) borrowed by some other group(s), the other not. One can roughly visualize this as follows:

```
--------------------------------------------------
! *t                                             !
! *k        ------------------------------       !
!           ! *?     ------------------  !       !
!           ! *k     ! *?             !  !       !
!           !        ! *?             !  !       !
!           !        ------------------  !       !
!           !                            !       !
--------------------------------------------------
```

Goddard (1978) makes a further argument against dialect mixing in Cheyenne: he observes that there is semantic specialization of the doublets in many cases (the glottal variant having non-diminutive meaning). He then gives examples of etymologically incorrect stop variants produced analogically from their glottal counterparts (with the more common hk for the less common hp). If I understand this puzzling argument correctly, he is trying to say that Ch hk is produced by a diminutivization rule, and hence not by dialect mixing. But this is a non sequitur: the synchronic production of Ch hk from Ch V? tells us nothing of the ultimate origin of such pairs (which he doesn't try to explain). Moreover, his rule doesn't even begin to account for Ch hp. Why would a diminutivization rule produce hp from V? in one case, hk from the same source in others - nearly always with the etymologically correct reflex of PA *p and *k?

In reality, semantic specialization is a very normal result of dialect mixing. In particular, one may expect innovating forms from prestige dialect(s) to be used in relatively formal situations, calling for respect, and their traditional counterparts in less formal ones. For example, while the cluster **Tl generally gives PA *t and W l (both from **tl) and Yurok r (from **?l), there are two kinds of exceptions. First, **-o:tetkohTliw-, **-o:tetkohCriw- 'kidney' 314 has W ?r rather than l (or its grade 2 variant r) as the reflex of **Cr - showing that the source is **?r rather than **cr. A more interesting example is **aTlemliyo:lkw- 'tell an origin myth' 115, and its root **aTl- 'tell experiences' 116. Here again Wiyot has the same reflex, presumably because such tellings are relatively prestigious activities. Consider also the use of the glottal variant of the second person prefix in Wiyot (described above) when the allocated kin are senior consanguineals - and thus entitled to extra respect.

Second, a suffix pair on the names for the digits of the left hand, considered sinister (**-VTl-VKha, **-VCr-VKha, **-VC^r-VKha 'digit from five to ten' 109, 412) evidently has the stop variants of both elements in all of the languages, including Yurok. The Yurok reflexes of **Tl and **Cr in this case are respectively t and s. In this case, the low prestige of the left hand is presumably responsible. The correspondences are:

| | **tl, **cr | | **?l, **?r | |
|---|---|---|---|---|
| PA | *t | *s | *l | *ŝ |
| Wiyot | l/r | ŝ | ?l | (?ŝ ?) |
| Yurok | t | s | r/l | (s ?) |

Hence, there is nothing at all surprising about doublets having specialized uses such as diminutivization both in Cheyenne and Yurok (_we?yon_ 'girl', but _we?yono?_ and, with the diminutive grade of vowel, _wɨ?yɨnɨk_ 'girls').

Goddard (1978) further argues against one specific theory of dialect mixing in Cheyenne (the Sutaio hypothesis) - but this is irrelevant for our purposes. So too is his discussion of gender dialects (_ke_ for women, _ĉe_ for men).

It isn't possible at present to be sure of what kind of dialects got mixed in Proto-Algic. They could have been based on geographical units (as in the better-known Indo-European cases, and also Montagnais [see Clarke 1981, 1988]), on lineage membership (as in Australia), or indeed on any type of grouping of which there were three or more in the society (one each for elements with pairs of stops, pairs of glottals, and _?_+_k_ elements).

## 8. Implications for Grammar

It's generally expected that when people adopt features of a prestige dialect, they will tend to primarily adopt highly visible ones. That is, features which will show up quickly and noticeably whenever they speak. High frequency verbs, pronouns, and grammatical elements are thus likely candidates. We have already seen examples of this in the present instance, such as the personal pronouns and prefixes.

Other cases, involving derivational and inflectional suffixes, where glottal variants are preserved are **-_eTew_ 'reciprocal' 193, 368a; **-_eCew_ 'reflexive' 192, 368b; **-_VTw_ or **-_VTew_ 'by tool, instrument, medium' 187, 369; and **-_Vŝ-eKoK_, **-_Vs-aKoK_ 'I-you pl.' 384.

We may now add several elements omitted earlier because they are too short to be good evidence for the existence of doublets.

Among the personal suffixes, there is:

(396a) *-T 'third person subject' (PA *-t and Y -?, Goddard 1991:65).

(396b) *-VK 'third person subject' 293, 300: PA *-k A499-500, W -ik (in verbs of being [somewhere]); W -i? (in relative clauses), Y -o?.

(397) **-VT 'second person subject' 212: PA *-at 'thou-her' A184, W -Vt T71 77, Y -a? 'thou-me' R70-72 (e.g., nekcena? 'you sg. meet me', ko?moyopa? 'you sg. hear me').

(398) **-VKw 'second person plural of verb': PA *-a:kw (in transitive *-eła:kw 'he...you', with *-eł 'second person object'), Y -o?w (for o-class verbs); PA *-e:kw (in all other combinations), Y -u? (for e-class verbs, presumably Pre-Yurok *-e?w). Compare Y -a?w for a-class verbs. The Algonquian endings are all conjunct order ones (see Bloomfield 1946:sec.46, 48); for the Yurok endings, see Robins (1958:33, 44).

There is one verb root:

(399) **T- 'be, exist, dwell' (Berman 1984), 221-222: PA *t(a:)- 'exist', W t(a)- 'stay, dwell, remain', Y ?(o)- 'be, exist, be born' and the near doublet Y to:?m- 'be together in a group' (with -o:?m 'incremental plural').

(400) **T- 'immobilizer: in one place', found in **Ta:wal- 'be around' 377 beside **a:wal- '(go) around' 056, and in **Tar-, **Tel- 'there' beside **er- 'thither' 255. Undoubtedly related to **T- 'be, exist, dwell'.

(401) **Tel-, **Tal- 'there' 223: PA *tał- A2023, W tɨlá- 'there, then, thus' beside **el-, **al- 'thither' 255.

(402) **T- 'WH-, TH-'. Surely related to **T- 'immobilizer: in one place' are **t- 253 and **?- 259 (the source for the first element in PA *ta:n- and *a:n- 'WH-') whose functions, while hard to reconstruct with precision, include locative, nominalizing, and relative-interrogative ones ('the one WH-/TH-'): W ĉi 'that's where, what, why, etc.', RhO a-bmi-noogseg 'the station' [i.e., 'where it (the train) stops going along'], e:-bngishmog 'in the west' [i.e., 'where it (heavenly body) sets'], e:-miijid 'what she eats', C e:-ki:-a:hkosit 'because she was sick' (Voorhis 1984a:38-1[9]); W ha-táłi?yɨk 'on my ship', halɨ-wɨ́mɨlutwu?y 'what one floats with' (Teeter 1964:82, 48); Y ?o ket?u?l 'there is a lake there', ?o tepo:noł 'in a forest', kic numi ?e?gah ?o lekwo?ł ku ?o?leł 'they were just eating when the house fell in', won so?n ?o ku yok ni hunowoni 'it is different from those that grow here'

(Robins 1958:103, 145, 146), W hi 'then (immediately after that)' [see Teeter 1964:88], Y ?i 'where, why, then'.

It is probably part of the normal evolution of zero-copula languages for locative pronominals to end up being used as verbs of 'being (somewhere)' and thence of existence (see Clark 1978 for the synchronic relationship between existential and locative constructions) - and I have suggested (Proulx 1991:155-157) that this is the origin of **T- 'be, dwell, exist'.

It is also normal for locative pronominals used as demonstrative pronouns to develop into nominalizers (Greenberg 1978), or into personal pronouns and thence into third person pronominal suffixes, and I have shown that this is a characteristic development in Algic (Proulx 1992:24-26). This explains the origin of *-T 'third person subject', and the similar nominalizer (W -?, Y -?) 299. The locative pronominal function of **T- must date back to Proto-Algic; probably the existential verb does too, but we cannot rule out independent parallel developments. However, it has several competitors for the functions of 'nominalizer' and of 'third person suffix', and these may well postdate the proto language. Compare **-VK 'third person subject' 396b, and:

(403) **-VK 'nominalizer': **-Vk 292 and W -V? 300, both 'nominalizer' - presumably from **k- 'the one previously mentioned' 251. An example of this nominalizer with Y -V? is **tepehtleKi 'ear' 092: W -tbɨlúk, Y -cpege?r (grade 2, infixed). After the loss of the preceding e, the glottal moved back onto the r.

Other likely doublets are:

(404) **-Ka 'demonstrative pronoun postclitic': **-ka (Algonquian, Wiyot, Yurok) and/or **-?a (Yurok, vowel hypothetical). This postclitic is generally found with **ya 'that, then' 243 (compare 240-241).

(405) **-VKh 'long thing' 155 evidently has a specialized form **-V? 'tree, stem' 308. The former is generally used with winding entities like rivers and ropes in Yurok and Algonquian, but **-VTl-VKha 'digit from five to ten' 109, 412 (lit. 'sinister finger, i.e., of the left hand') suggests that the broader meaning found in Wiyot is ancient.

## 9. Implications for the Status of Laryngealized Consonants

With those in the prefixes (and some others) shown to be secondary, there is no further evidence outside of Yurok for glottalized consonants, as Goddard (1991:65) points out. Outside of Wiyot, there has never been any evidence for aspirated stops, nor of glottal catch as the first member of a consonant cluster. Nor, outside of Algonquian, of PA *h as first member of a consonant cluster. In each of these cases, the features in question must be stated to have dropped in the other branches of the family before the rules of sound change apply in those branches.

There are often allomorphs with simple stops beside others with aspirated ones in Wiyot (Proulx 1985a:67); some Yurok suffixes have variants with glottalized and nonglottalized consonants (209, 213); and **s?yoyeTkw- 'slip' 86 has two grade variants in Yurok, one with its initial consonant glottalized, the other not: t?yoykwo?n 'it's slick' (with **-Vn abstract final 'nondeliberate action' 177, and **-T 'third person' 396a) and syoykweł 'a slippery place' (with **-Vł 'nominalizer' 298). In addition, the personal pronoun Y ke?l 'thou' lacks glottalization of the initial k, versus the prefix Y k?e- 'thy' (as pointed out by Goddard 1991:65). However, this may merely reflect the universal tendency for complex stops to have restricted distributions within words (see Proulx [1974] for examples from Quechua). In polysynthetic languages like the Algic ones, restrictions of this kind give rise to allomorphs which presumably may persist in many cases even after the original distributional rules break down.

The same is clearly true of W ? before consonants (see Teeter 1964:22-26 for the distributions). Preconsonantal h in Algonquian is more stable, but even it disappears unpredictably in some cases. In Menominee, for example, there is M pes- from *pehŝ- 'peel, husk' S153, and M seko:h 'weasel' for expected *sehko:h and M pes<:hnew 'she has a foriegn body in the eye' for expected *pehs<:hnew S131. Fox and Shawnee attest pseudo-PA *peŝekehsiwa 'deer' (F peŝekesiwa, Sh pŝekθi), an irregular diminutive of *peŝehkiwa 'buffalo' A1858. In addition, compare the root in PA *tahkon- 'grasp (i.e., grab with the hand)', having final **-Vn 'by hand' ǂ82, with the root in *sakipw- 'bite (i.e., grab with the teeth)', having final PA *-pw 'by mouth' (Bloomfield 1946:sec.84). The latter root, which is evidently a diminutive grade of the former, lacks its preconsonantal *h.

Restricted distributions or unpredictable losses are not necessarily indications of recent origin. Laryngeals are there in many words with Proto-Algic etymologies, and there is no obvious source for most of them but inheritance. It is not satisfactory to dismiss them as resulting from difusion (pace Goddard 1991:65), unless one can explain how they came to be in some words rather than others.[8] After all, words with Proto-Algic etymologies are rather unlikely to be post-Proto-Algic loans.

Examples contaianing Wiyot ? are: **?eki?yem- 'hat' 047, **a:?lewe 'projectile point' 315b, **i:?meli 'below' 124, **pe?meyi 'grease' 024, **wa:wa?lewi 'egg' 003. Examples with Yurok C? are: **as?ola:?w- 'pound' 120, **ek?ey- 'conceal' 122, **le:k?awi 'sand' 072, **mey?ełki 'nettle' 062, **s?yoyeTkw- 'slip' 86, 392. Examples with Wiyot Ch are: **meyehkhwel- 'weep, mourn' 089, **nikhl- 'three' 046, **Tehkhon- 'grasp, hold onto something with the hand' 064/019, **thigw- 'through an opening; out' 288, **wechowe 'quadruped-tail' 317, **weTa(:)phega 'root, fine ... used in sewing (usually spruce)' 102.

There is only one of these items for which I have a possible explanation: **as?ola:?w- 'pound'. The initial **a drops in Yurok, and Yurok words describing violent actions tend to have **s? rather than **s in initial position. Of the 15 words listed by Robins (1958:lexicon) with initial Y s?, 10 have such meanings. Only two words with such meanings ('break' and 'kick') have initial Y s. Compare also Y s?o:p- 'hit' (e.g., s?o:pe?weyet- 'hit in the face') with Y to:ps 'slap' (with transitive -s, and grade 1 alternation of s to t). All this suggests that at least some of the Yurok words in initial s describing violent actions may have acquired glottalization secondarily.

Moreover, we now have an explanation for the developments of **s?yoyeTkw- 'slip' 86, 392 in Yurok: Y syoykw(eł) 'a slippery place' (with **-Vł 'nominalizer') beside t?yoykw(o?n) 'it's slick' (with **-Vn abstract final 'nondeliberate action' 177, and **-T 'third person'). This case is the exact opposite of the 'hit' and 'slap' pair - and for good reason: 'slipping' and 'being slick' are in themselves nonviolent actions. Evidently, just as s-initial Yurok words have tended to add glottalization where violent action is involved, they have tended to suppress it otherwise. However, in this case we are explaining the LOSS of **? in Yurok, not its origin. More importantly, words with initial Y s are an isolated case which cannot be extended to the language as a whole.

Two questions arise in the present circumstance, which it's helpful to keep separate. (1) What is the phonemic status of the features in question in the proto language? (2) How should proto words be transcribed? For there is no absolute law requiring that transcriptions be strictly phonemic (provided contrasts are all recoverable from them) - as Goddard (1979:74-75) has pointed out.

In the present case, the existence of complex stops in Proto-Algic is uncertain; they may in fact all be secondary. But there are very good reasons to write them anyway. If one transcribes the complex stops in the proto words, one doesn't so easily forget about them. Hence, one may more easily spot ways in which they may have conditioned changes in one of the other branches. This has not happened to date in the present case, but one cannot for that reason entirely rule out the possibility that it yet may. Also, a transcription including complex stops may be more helpful in wider comparisons with other language families: more distant relatives, if identified, might yet provide an explanation for them.

Finally, transcribing them helps keep us alert for possible sources for them, which may sometimes carry grammatical information. For example, **-Vp? 'place, put in place' draws attention to the sequence PA *a?t- A171 (rather than the root PA *a?- A159, the obvious comparandum for **-Vp), and thus leads to the reconstruction of the sequence:

(406) **-Vp-eT 'put SOMETHING in place' 161, 178: PA *-a?t 'place it', Y -ip? 'place things'. (The second element is **-VT 'nonpersonal object' 178).

This new reconstruction suggests that Proto-Algic had a transitive formative where Yurok does not, at least in this case. Another example involves the comparing of Y kep?- 'block an opening' with the PA sequence *kepah- 'block it' A734-739 (with *-ah 'by tool, instrument, or medium' A76) rather than root *kep- 'block' A728, as follows:

(407) **kep-e?- 'close an opening': PA *kepah- A734-739 (root PA *kep- A728-742, 744-745, 747-757, and a-grade *-ah 'by tool, instrument, or medium' A76), Y kep?- (in kep?eɬ 'be deaf', kep?oɬ 'there's a barrier', kep?oɬ nɨp?ɨ?n 'my nose is blocked', kep?oksine?m 'you plug it up [e.g., a rat hole]'). Cf. V cappetaw 'deaf', and PA *kakye:pehte:wa redupl. 'she's deaf' S66.

(408) **Kep-etkoy?- 'choke on something (food, tobacco smoke)': fwC kipiskóyoo 'she chokes eating', Y ?epkɨ?y- 'choke smoking'. The root **Kep- is also seen in PA *kepene:we:ne:wa 'she chokes him' A745, *kepaɬa:mowa 'she chokes' A742, Y ?ep- 'be choked'. Presumably, *t drops between obstruents in Yurok.

(409) **kep-w- 'cover': PA *kepw- A756-759, W kwɨp- (with distant metathesis). The stem extension is **-Vw 357. For a similar metathesis in Algonquian, see A726 'northern pike' and Middle Atlantic Algonquian *kwen- 'long' from PA *kenw- (Goddard 1980:148).

Similarly, **VT can be discerned behind Wiyot aspiration in W hɨph- 'tie', helping to establish the identity of a root, medial, and final:

(410) **apiT-, **epiT- 'tie, cord' 347. The loss of **i in Wiyot is evidently regular in this environment, compare W bɨpt 'tooth' with PA *mi:pit- and medial *-a:pit-.

**apiTeTk(w)-, **epiTeTk(w)- 'tie into a round shape, i.e., a bunch or bundle' 346, 389.

(411) **apiTe:kh-, **-a:piTe:kh 'cord' 348: Y pekcic [diminutive?] 'thread, string, rope', dependent -pek; medial PA *-a:pye:k 'string' A59. Composed of **-VpiT 'tie, string, root' 162, and **-Vkh 'long thing' 155. Presumably, **iT ---> **i? ---> *i ---> PA *y. Yurok has no initial p? in the lexicon, so presumably **piT ---> **pe? ---> **p? ---> p. The dependent stem would be analogical.

(412) **-VpiT- 'tie, string, root' 162: PA *epit- as in *kaškepit- 'tie it shut' A614-615, PA *-a:py (as in NiO mitigwaabiig 'bows' beside mitigoon 'sticks' from Pre-O *mehtekw-a:py-aki and *mehtekw-ali), Y -pet as in sekipetek? 'I tie it securely in place'.

Details of this sort are much easier to spot if one is transcribing Proto-Algic reconstructions with the complex stops their daughter languages suggest. Out of sight is out of mind. Moreover, it is a classical logical fallacy to suggest that because some of the complex stops of Wiyot and Yurok are secondary, all of them are. One of the most universal of phonological developments is for segments to lose some of their articulatory features, being reduced to simpler elements like glottal catches and aspirations. Such simplifications may in principle be repeated many times in the history of a language, with accretions to simpler phonemes comming from various sources at different times. At the same time, the simplest phonological elements are the weakest and most unstable, commonly being restricted in their distribution (e.g., being limited to one per word) or simply disappearing in enviroments where they are less prominent. To dismiss the evidence of earlier structure which they provide, as one does in too quickly devaluing and dismissing them as "secondary" and "unstable", is simply to throw out an important source of data which, at Proto-Algic time depths, is all too scanty to begin with.

## NOTES

1. Languages, their abbreviations, and the sources from which they are generally cited are as follows: Abenaki-Ab-Laurent (1884), Day (1964); Plains Cree-C-Bloomfield (ms.); Swampy Cree-swC-Voorhis (1984a); Western Cree-fwC-Faries and Watkins (1938); Delaware-D-Goddard (1969)<uD=Unami, mD=Munsee>; Fox-F-Bloomfield (ms.); Kickapoo-K-Voorhis (1974); Loup-L-Day (1975); Mahican-Mh-Mastay (1982); Menominee-M-Bloomfield (1975); Miami-Mi-Voegelin (1937-40); Micmac-Mc-Proulx (field notes), DeBlois and Metallic (1984); Natick-N-Trumbull (1903); Ojibwa-O-Bloomfield (1957); Central Ojibwa-bO-Baraga (1878); Western Ojibwa-NiO-Nichols (1979); Central and Eastern Ojibwa-RhO-Rhodes (1985); Passamaquoddy-Ps-LeSourd (1984); Penobscot-Pe-Voorhis (1979); Proto-Algic-PAc-Proulx (1984); Proto-Algonquian-PA-Aubin (1975), Siebert (1975); Saulteaux-wO-Voorhis (1984b); Shawnee-Sh-Voegelin (1937-40); Wiyot-W-Teeter (1964); Yurok-Y-Robins (1958), Proulx (field notes).

PA reconstructions found in Aubin (1975), and Siebert (1975) are respectively identified with the letters A, and S plus the item number. The frequent citations from Bloomfield (1946), Robins (1958), and Teeter (1964) are respectively identified with the letters B, R, and T plus the page number. Citations of Wiyot texts (Teeter 1964) are followed by the text number and line; those from my Yurok fieldnotes are identified by NB plus notebook number and page.

Transcription generally follows that of Siebert (1975) for Algonquian, Teeter (1964) for Wiyot, and Robins (1958) for Yurok. However, the following changes have been made: PA *_ɫ_ is written for *_θ_, PA *_s_ for cedilla, PA *_t_ for *_x_, W _?_ for _h_ before a consonant, W _a_ for _o_, W _ɨ_ for _a_, and Y _ɨ_ for inverted _r_. For discussion of the changes, see Proulx (1984:168-169). Orthographic concessions to my word processor: _s_ wedge is written as _ŝ_, _c_ wedge as _ĉ_, schwa as _ɨ_, and Menominee epsilon as _<_.

2. Nevertheless, Berman (1984:fn.6, 7, 8, 9, 12) does indicate some specific differences with my 1984 reconstructions - and these I can certainly address. In his fn.6, he objects to **_a:wal_- 'around' 056: Mc _al_-, W _hal_-, Y _ho:r_-. He evidently wants to explain 'around' as secondary from **_a:_- 'go' 220 (see Berman 1984:336), but there are problems. First, he has no explanation for the _l_ in Micmac nor _r_ in Yurok. Second, the proposed direction of semantic drift is wrong. Compare the following Micmac, Yurok, and Wiyot verbs (DeBlois and Metallic 1984, Proulx 1985b, Teeter

mss) as to root and meaning:

(1) Mc *ala:sit* 'she goes about', Y *ho:le?m* '(they) are around, go, travel, fare', W *halú?wi?* 'boat, (that which) goes around'.

(2) Mc *alsink* 'she flies around', Y *ho:letkoli?m* 'birds circle in flight'. Compare Mc *pemsink* 'she flies by in the air after someone hits her', and Y *la:yetkoli?m* 'they all fly by'.

(3) Mc *ala:x* 'she swims about', Y *ho:ru:rek?* 'I swim around', W *ŝu?r* *halúliɬ* 'sea serpent' (means 'swims around in the ocean'). Compare, Mc *pema:x* 'she swims along', Y *ra:yurek?* 'I swim (along)', and W *hutúliɬ* 'she's swimming towards me'.

(4) Mc *ala:lukwet* 'she floats about', Y *ho:lenek?w* 'it drifts around', W *halɨwɨ́mɨlutwu?y* 'what you float around with'. Compare W *wɨ́mɨlutwuy* 'they (indef.) float with it' T48, and Y *la:yonek?w* 'it drifts by'.

(5) Mc *alu:lɨk* 'she rows him about', Y *ho:lecok?* 'I move my boat around with a paddle'. Compare Y *ra:yecok?* 'I paddle along'.

(6) Mc *ala?ulat'm* 'she carries it around on her back', Y *ho:lu:lesek?* 'I carry a pack', Y *ho:lewkwse?m* 'you have a load in your boat', W *hálulɨwiɬ* 'she brings it along, carries it around'. Compare rMc *enma?ulat'm* 'she carries him home', Y *la:yu:lesek?* 'I come by here with a pack', and Y *kelomewkwse?m* 'you turn around and come back in your boat', and W *ta* *tɨ́kwulɨwuy* 'they bring it down', with *tikw*- 'down' T32.

(7) Mc *al'ĉa:t* 'she staggers, stumbles about', Y *ho:letok?* 'I walk with effort (in a steep place, under a heavy load)'. Also Ps *alálikɨnɨ́ke* 'she walks around in deep snow (without snowshoes)'. Compare Mc *pem'ĉa:t* 'she leaps or hops along', Y *himetos* 'hurry!'.

(8) Mc *aliskalk* 'she feels around for him', Y *ho:letewek?* 'I grope around'. Compare Y *la:yecewek?* 'I grope my way along'.

(9) Mc *alo:stasit* 'she hints', Y *ho:roksek?* 'I have clever but changing and unreliable thoughts'. Compare Y *ca:noksek?* 'I forgive (lit., have new thoughts)'.

Consider also: Mc *alamk* 'she looks around for him', Mc *al'kopiĉik* 'they all sit about', and Y *ho:lecok?* 'I stir food with a paddle cooking'. For the Micmac segmentation, see Inglis (1986, notably pp.133-134).

It seems evident to me that core meaning of the root is 'an aggregation of individual actions (or discrete portions of an action), not coordinated with each other, but all of the same type'. Typical is 'moving about', with apparently random changes of direction. With semantic bleaching, 'going about' can become 'going, traveling' (for example, in Coyote stories) - but the reverse path of semantic evolution is unthinkable. There is just no way from 'I go' to 'I'm stirring food with a paddle cooking', or many of the other meanings.

Berman also claims that the vowel length in Y ho:l- is secondary, and sends us on a fruitless paper chase for the evidence: first to Berman (1982a:417), where he admits that 'length, though, does not occur as a laryngeal increment with the same regularity as h or ?', and then to Berman (1981:257-258), where the discussion is only about h and ?, not vowel length in Yurok. The only discussion of vowel length is of the Proto-Pomo canonical form CVHCV̸, where H can be either (unpredictably) vowel length or (predictably) h or ?.

All this is presented as an alternative to accepting the straightforward reconstruction of **a:wa for PA *a:wa (Mc a, lengthened in monosyllables), Y o:, and W a. This despite such cognate sets as PA *wa:walwi 'egg' A2135, Y wo:lew 'roe of several fish', and W wá?l 'salmon roe' 003; and PA *na:wal(a)w- and Y no:?rep- 'following one behind the other' 055 (where PA adds postmedial *-(a)w, and Yurok the intransitive final -ep). PA *a:wa generally gives Mc a, as in Mc saxamax 'chiefs' from PA *sa:kima:waki.

He also questions the etymology of Wiyot hálikwɬiɬ 'grey fox (the one who weeps around)'. Well, it really does consist of W hal- 'around' and -ikwɬ 'weep', as is evident from ta pɨ́kwɬiɬ 'she's crying secretly' (with p- 'secretly', see Teeter 1964:52), kawɨrɨwɨ́kwɬiɬ 'they all start to cry' (with kaw- 'begin' and -ɨrɨw 'all', see Teeter 1964:51-52). Especially after a preverb ending in a vowel, Wiyot often inserts an empty root l- before a medial beginning in another vowel (Proulx 1985a:66). This ambiguity of l- led Teeter (1964:85) to synchronically analyze hal- as a sequence of preverb ha- 'around' plus empty l-, but this is historically incorrect and evidently misled Berman.

Berman's fn.7 questions my direct equation of the roots in PA *no:n-, W dunaĉ-, and Y newon- 'suck at the breast' 006 - claiming that the Yurok form has initial change (of the Ojibwa type, **o: to **wa:). This was a plausible enough hypothesis when he formulated it, but no further evidence for this type of change in Proto-Algic has since been found. Rather, what we find is ew(o) = u in such Yurok sets as no:lum- 'love' and passive no:lew (uninflected verb) and no:lewomoy-. These show a Pre-Yurok contraction of *ewo to *o: outside a first syllable except where blocked by word shortening (and thence analogically restored). The original environment was probably 'in unstressed position', as is the case for the contraction of W ɨwu to u (Teeter 1964:26). This suggests we may be dealing with a very low level phenomenon here.

In his fn.8, he scoffs at my "unique" *o: + ɨ + o correspondence in **kemotl- 'steal' 21 - ignoring the fact that the vowel length in Algonquian is stated to be secondary to the simplification of the cluster **tl, where added vowel length is regular (Proulx 1984:193). PA *o is a rare phoneme (see Goddard 1979:75, who even doubts its existence), so there should be no surprise if we only have one example of the (entirely regular) *o + ɨ + o correspondence modified by the compensatory lengthening for the loss of **l in PA (the Wiyot and Yurok reflexes are also found in 049 and 120). The distribution of **o in Proto-Algic is limited to "next to **l or **tl (and usually **m) in the second syllable of a verb stem whose first syllable has a short vowel and is open" - which may mean it was a recent innovation in the protolanguage, but does not invalidate the correspondence.

He also (fn.9) scoffs at my allegedly "special vowel *i" in **ihkwa 'louse' 009, supposedly found only in 3 items before **k. But he is looking only at Algonquian + Wiyot pairs, ignoring the important Yurok cognates. **i (*e + i + a) was found in 009, 032, 046, 047, 099, 111 (before **t) in my 1964 paper, and since has been found in 385. There is nothing special about it.

Finally, in his fn.12, he questions my reconstruction of **ne:w- 'see' 043 (PA *ne:w-, Y new-) on the grounds that some inflectional suffixes of this verb in Yurok optionally have long link vowels generally found only after roots lacking a vowel (i.e., in monosyllables). His idea is that the Yurok root must therefore once have lacked a vowel (**nw-). However, vowels in monosyllables do not generally lengthen in Yurok (witness tey 'brother-in-law' and Y ł?os 'take!', with root ł- 'take' and imperative singular -?os) - and long vowels in inflectional suffixes of this sort are also found outside of monosyllables (e.g., merkweta:k? 'I ate everything' NB 4:38 and skewip?a:?m 'you put in order' R44). So we are dealing with a morphological class with archaic irregularities and not with a phonological environment. Small irregular morphological classes tend to be shrinking ones, and there is nothing odd about leftovers like Y new- and merkwet- remaining in a class which otherwise has come to be limited to zero-vowel roots.

3. In the case of nouns, in Wiyot the preceding **K of the second person prefix produced aspiration in a following obstruent and the preceding **w of the third person prefix metathesized with a following **k - and so was preserved there, though lost before

other consonants (see Goddard 1966). The preceding **n of the first person prefix was lost, but despite this the prefix was not restored in some nouns - notably a class of kin nouns. They were restored elsewhere, including with subordinative verbs, in both Wiyot and Yurok. In Proto-Algic, nouns referring to parts of wholes generally took a third person prefix (allocating the part to the whole, see Proulx 1992:sec.5.1). These prefixes are generally restored as well, unless a semantic shift has ended the part-of-whole relationship (as in the case of Y skoy 'strip of buckskin' - beside PA wa?lakaya [a-grade] and W wɨ́tkay 'skin' 280).

In the case of verbs, restoration of an initial stem syllable was possible by analogy with shielded forms (prefixed ones, those with preverbs, or infixed ones, where the **g of the infix **-eg would block the rule). Thus we have Y nek(e?y) 'call by name' beside PA *ni:hk- (Pre-PA *negehk-, with contraction of the infix **-eg plus the following **e to PA *i:, see Proulx 1984:197). Shielded forms are rare for verbs used as adjectives, making restoration unlikely: Y pek(oy)- 'be red' beside PA *nepek- and *ni:pek- 'be bloody' 125, and W ɬáw(ik) 'be distant' beside PA *wa:?law- (with archaic ablaut) 279.

4. However, **nepi?i 'water' 067 (Y pa?ah, PA *nepyi) suggests it may be lost between like vowels in a word final sequence in PA - or perhaps after front vowels (**i?i > *iy?i > *iyi > PA *yi). This would make PA *-ihi 'obviative plural' analogical from its byform *-ahi, on the model of *-ili and *-ali 'obviative singular', as the expected reflex of Pre-PA **-i?i would be *-yi.

5. Although it does so in such pairs as W wɨtw 'heart' and wɨtwi?l 'her heart', it's unlikely that obviation disambiguated these meanings in the proto language (see Proulx 1991:142-143).

6. There are some apparent examples of an initial **t preserved in verb medials in Proto-Algic, but there is some evidence that the **t was not element-initial except after the loss of a previous syllable. For example, beside **-tekwli 'heart' 112 and medial **-Vtekwl- 'think' (see 416 in appendix B below) there is W dɨtw- 'have thoughts, think' (Reichard 1925:127, in normalized orthography). This suggests a verb stem **netekwl- from which both the medial and the dependent noun are formed.

7. Of course, single-element variants must be consistent as to having stop or glottal reflexes. An isolated apparent exception to this rule is **weTegeteke, **weTecake 'her buttocks' 327 - which apparently contains both **T and **t (and its grade 2 variant). However, Y -cɨk is easily explained as resulting from reinterpretation of stem-initial **t as part of the prefix, and

the regular analogical replacement of the stop variant of the prefix by its glottal variant (and thus Y _?w_- 'her', see 367).

8. Such explanations are not impossible; I have explained the origin of aspirated and glottalized stops in Quechua (Proulx 1974), but I find no such explanation in the Algic case.

## REFERENCES

Aubin, George F. 1975. _A Proto-Algonquian dictionary_. National museum of man Mercury series, Canadian ethnology service paper no. 29. Ottawa: national museums of Canada.

Baraga, R.R. Bishop. 1878. _A dictionary of the Otchipwe language_. Reprinted by Ross and Haines. Minneapolis, Minnesota, 1973.

Berman, Howard. 1981. Review of The Languages of Native America: Historical and Comparative Assessment (Campbell and Mithun). _IJAL_ 47.248-262.

Berman, Howard. 1982a. Two phonological innovations in Ritwan. _IJAL_ 48.412-420.

Berman, Howard. 1982b. A supplement to Robin's Yurok-English lexicon. _IJAL_ 48.197-222.

Berman, Howard. 1984. Proto-Algonquian-Ritwan verbal roots. _IJAL_ 50.335-342.

Berman, Howard. 1986. A note on the Yurok diminutive. _IJAL_ 52.419-421.

Berman, Howard. 1990. New Algonquian-Ritwan Cognate Sets. _IJAL_ 56(3):431-434.

Bloomfield, Leonard. 1946. Algonquian. _Linguistic structures of native America_. Ed. Harry Hoijer, pp. 85-129. New York: Viking fund.

Bloomfield, Leonard. 1957. Eastern Ojibwa. Ann Arbor: University of Michigan press.

Bloomfield, Leonard. 1975. Menominee lexicon. Ed. Charles F. Hockett. Milwaukee: Milwaukee public museum.

Bloomfield, Leonard. Ms. Fox and Cree dictionaries. Photocopies distributed by C.F. Hockett.

Bybee, Joan L, William Pagliuca, and Revere D. Perkins. 1990. On the asymmetries in the affixation of grammatical material. Studies in Typology and Diachrony (for Joseph H. Greenberg). Croft, William, Keith Denning, and Suzanne Kemmer, eds. Typological studies in language 20. Amsterdam - Philadelphia: John Benjamins (1-42).

Clark, Eve V. 1978. Locationals: Existential, Locative, and Possessive Constructions. Universals of Human Language, vol.4 (Syntax). Ed. Joseph H. Greenberg. Stanford, California: Stanford University Press (83-126).

Clarke, Sandra. 1987. Direct Convergence in a Non-Overtly Stratified Society. Proceedings of the Fourteenth International Congress of Linguists. Eds. Werner Bahner, Joachim Schildt, and Dieter Viehweger. Berlin: Akademie-Verlag (1335-1339).

Clarke, Sandra. 1988. Linguistic Variation in the Non-Stratified Social Context. Ed. Alan R. Thomas. Methods in Dialectology. Clevedon: Multilingual Matters (684-699).

Day, Gordon M. 1964. A St. Francis Abenaki vocabulary. IJAL 30.371-92.

Day, Gordon M. 1975. The Mots Loup of Father Mathevet. Ottawa: National museums of Canada publications in ethnology 8.

DeBlois, Albert D. and Alphonse Metallic. 1984. Micmac Lexicon. Canadian ethnology service paper no. 91. Mercury Series, National Museum of Man. Ottawa.

Eliot, John. 1666. The Indian Grammar Begun. See Trumbull.

Ellis, C. Douglas. 1971. Cree verb paradigms. IJAL 37.76-95.

Faries, R. and E.A. Watkins. 1981[1938 + 1865]. Dictionary of the Cree language. Toronto:Anglican Book Centre.

Glenmore, Josephine Stands in Timbre and Wayne Leman. 1984. Cheyenne Topical Dictionary. Cheyenne Translation Project.

Busby, Montana.

Goddard, Ives. 1966. Review of The Wiyot Language by Karl V. Teeter. *IJAL* 32.398-404.

Goddard, Ives. 1969. *Delaware verbal morphology*. Harvard university doctoral dissertation.

Goddard, Ives. 1974. An outline of the historical phonology of Arapaho and Atsina. *IJAL* 40.102-16.

Goddard, Ives. 1978. The Sutaio dialect of Cheyenne: A Discussion of the Evidence. *Papers of the Ninth Algonquian conference*. Ed. William Cowan. Ottawa: Carleton university (66-80).

Goddard, Ives. 1979. Comparative Algonquian. *The languages of native America: historical and comparative assessment*, ed. Lyle Campbell and Marianne Mithun. Austin: university of Texas press (70-132).

Goddard, Ives. 1980. Eastern Algonquian as a Genetic Subgrouping. *Papers of the Eleventh Algonquian conference*. Ed. William Cowan. Ottawa: Carleton university (143-158).

Goddard, Ives. 1988. Pre-Cheyenne *y. *In honor of Mary Haas: from the Haas Festival Conference on Native American Linguistics*. Bill Shipley. Berlin: Mouton de Gruyter, 345-360.

Goddard, Ives. 1991. Algonquian Linguistic Change and Reconstruction. *Patterns of Change, Change of Patterns: linguistic change and reconstruction methodology*. Ed. Philip Baldi. New York - Berlin: Mouton de Gruyter.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? *Universals of human language*, vol. 3 (word structure). Stanford, California: Stanford university press.

Haas, Mary R. 1970. Consonant symbolism in northwestern California: a problem in diffusion. *Languages and cultures of western North America*. Earl H. Swanson, Jr. Pocatello: Idaho state university press.

Hamp, Eric P. 1970. Wiyot and Yurok correspondences. *Languages and cultures of western North America*. Earl H. Swanson, Jr. Pocatello: Idaho state university press.

Hockett, Charles F. 1948. PotawatomiI-III. *IJAL* 14.1-10, 63-73, 139-149.

Inglis, Stephanie Heather. 1986. *The fundamentals of Micmac word formation*. M.A. thesis, Memorial university of Newfoundland.

Laurent, Joseph. 1884. *New familiar Abenakis and English dialogues*. Quebéc: Leger Brousseau.

Leavitt and Francis, eds. 1986. *Kolusuwakonol peskotomuhkato-wolastoqewi naka ikolisomani latuwewakon/ Philip S. LeSourd's Passamaquoddy-Maliseet and English Dictionary*. Fredericton, N.B.: University of New Brunswick (Micmac-Maliseet Institute).

LeBoulanger, Jean-Baptiste. 1725. [French-Illinois dictionary]. (Manuscript in the John Carter Brown library, Providence, R.I.)

Leman, Wayne and Josephine Stands in Timber Glenmore. 1984. *Cheyenne Topical Dictionary*. Busby, Montana: Cheyenne translation project.

LeSourd, Philip. 1984. *Kolusuwakonol*. Ed. Robert M. Leavitt and David A. Francis. Fredericton, N.B.: University of New Brunswick (Micmac-Maliseet Institute).

Masthay, Carl. 1982. *Schmick's Mahican manuscript, transcribed and rearranged by English translation*. St. Louis: published by the author.

Mathevet. 1748. See Day, 1975.

Nichols, John and Nyholm, E. 1979. *An Ojibwa Word Resource book*. St. Paul: Minnesota archaeological society.

Piggott, G. L., and A Grafstein. 1983. *An Ojibwa lexicon*. Canadian ethnology service paper no. 90, Mercury series. Ottawa: National Museum of Man.

Proulx, Paul. 1974. Certain aspirated stops in Quechua. *IJAL* 40.257-262.

Proulx, Paul. 1982a. The linguistic evidence on the Algonquian-Iroquoian encounter. *Approaches to Algonquian Archaeology* (Proceedings of the thirteenth annual conference of the archaeological association of the university of Calgary, 1980), pp. 189-211.

Proulx, Paul. 1982b. Yurok retroflection and vowel symbolism in Proto-Algic. *KWPL* 7.119-123.

Proulx, Paul. 1984. Proto-Algic I: phonological sketch. *IJAL*

50.165-207.

Proulx, Paul. 1985a. Proto-Algic II: verbs. IJAL 51.59-94.

Proulx, Paul. 1985b. Notes on Yurok derivation. KWPL 10(2):101-144.

Proulx, Paul. 1988. The demonstrative pronouns of Proto-Algonquian. IJAL 54:309-330.

Proulx, Paul. 1990. Proto-Algonquian verb inflection. Studies in Native American Languages VI. KWPL 15(2).100-145.

Proulx, Paul. 1991. Proto-Algic III: pronouns. KWPL 16.129-170.

Proulx, Paul. 1992. Proto-Algic IV: nouns. KWPL 17.11-57.

Rand, Silas Tertulius. 1971 [1894]. Legends of the Micmacs. Johnson Reprint Corporation.

Reichard, Gladys A. 1925. Wiyot grammar and texts. University of California publications in American archaeology and ethnology 22.1-215.

Rhodes, Richard A. 1985. Eastern Ojibwa-Chippewa-Ottawa dictionary. Ed. Werner Winter. Trends in linguistics: documentation 3. Berlin - New York -Amsterdam: Mouton.

Robins, R. H. 1958. The Yurok language. University of California publications in linguistics 15, Berkeley.

Siebert, Frank T, Jr. 1967. The original home of the Proto-Algonquian people. Contributions to Anthropology I: Linguistics I (Algonquian). Bull. 214, anthropological series 78. Ottawa: National Museums of Man.

Siebert, Frank T, Jr. 1975. Resurrecting Virginia Algonquian from the dead. Studies in Southeastern Indian languages. Ed. James M. Crawford, pp. 285-453. Athens: university of Georgia press.

Swadesh, Morris. 1971. The origin and diversification of language. Chicago: Aldine.

Taylor, Allan. 1969. A grammar of Blackfoot. Universtiy of California at Berkeley doctoral dissertation.

Teeter, Karl V. 1964. The Wiyot language. UCPL 37. Berkeley and Los Angeles: university of California press.

Teeter, Karl V. 1971. The main features of Malecite-Passamaquoddy grammar. _Studies in American Indian languages_. Ed. Jesse Sawyer, pp. 191-249. UCPL 65. Berkeley and Los Angeles: university of California press.

Teeter, Karl V. Mss. _Wiyot dictionary_. In possession of the author.

Trumbull, James H. 1903. _Natick dictionary_. BAE-B 25. Washington, D.C.

Turner, David H. and Paul Wertman. 1977. _Shamattawa: the structure of social relations in a northern Algonkian band_. Canadian ethnology service paper no. 36, Mercury series. Ottawa: National Museum of Man.

Voegelin, C.F. 1937-40. Shawnee stems and the Jacob P. Dunn Miami dictionary. _Indiana Historical Society_. Prehistory research series 1:63-108, 135-167, 289-341, 345-406, 409-478.

Voorhis, Paul H. 1974. _Introduction to the Kickapoo language_. Language science monographs 13. Bloomington: Indiana university.

Voorhis, Paul H. 1979. _Grammatical notes on the Penobscot language from Frank Speck's Penobscot transformer tales_. University of Manitoba anthropology papers 24. Winnipeg: university of Manitoba.

Voorhis, Paul H. 1984a. _A Cree phrase book_. Brandon university department of Native Studies.

Voorhis, Paul H. 1984b. _A Saulteaux (Ojibwa) phrase book based on the dialects of Manitoba_. Brandon university department of Native Studies.

## Appendix A: NEW OR SIGNIFICANTLY REVISED RECONSTRUCTIONS

Some of these reconstructions are among the most problematic ones in Proto-Algic, and are included here only for the sake of completeness.

(413) **_weTayehca_ (formally diminutive) '(her) dog': PA *_wetayehsa_ 'her dog' (M _oti:hsak_ 'her pet dogs') and W _wáyic_ 'dog', and Y _wɨyc?_(_ɨks_) 'dog', adding -_Vks_ 'child' (as in Y _ca:nu:ks_ 'baby', with _ca:n_- 'new'). Evidently, there was early

metathesis of the first two vowels in Pre-Yurok, giving *wa?eyehca. With the loss of **h, the two **e's and the final vowel, *wa?yc fed into the glottalization rule. The only real difficulty is that a pet dog is not literally part of a whole - though, at least in Algonquian, the noun is clearly dependent (as would be appropriate for a body part).

(414) **wa:yelKwa 'roe': PA *wa:hkwa S168, Y ?wɨył 'egg'. The Yurok term evolves via Pre-Y *woył?wɨ (and, with vowel harmony, *wɨył?wɨ) followed by metathesis of *?wɨ with *wɨ and loss of the latter in wordfinal position.

(415) **-VTl-VKha, **-VCr-VKha, **-VCˆr-VKha 'digit from five to ten' 109: (a) W -ɨlúk in tɨkłɨlúk 'six', dɨlúk 'ten', and (with consonant grade harmony, T21) bɨŝɨrúk 'nine'; (b) Y -ɨsik? in cɨwɨsik? 'seven', cf. cɨwɨhs- 'point', and grade 1 Y -etik? in knewetik? 'eight', with knew- 'long'; (c) PA *-a:ŝika, as in *nekwetwa:ŝika 'six' A1048 (with **nekwet- 'one' 63, both originally reconstructed with *o for *we). The literal meaning is 'sinister long things', in reference to the fingers of the left hand (presumably counted on). Possibly the rounding in Wiyot was caused by a plus of **w in stem final position. This and the secondary glottalization in Yurok could respectively be the nominalizing suffixes **-Vw 304 and **-T 299, the former also present in the PA suffix *-ikwi 'digit from five to ten' (Siebert 1975:303, reconstructed with *ye for *i). See item 155 on the reconstruction of **-VKha.

(416) **we?ilkw-, **we?irkw- 'bundle-strap' 037: (a) PA *wi:hkwe:- 'wrap with a bundle strap' (with final *-e:), Y ?weskul 'strap' (with the nominalizer **-Vl 302); (b) PA *wi:ŝkwe:- 'wrap with a bundle strap'. **? may well be from **T, but a stop variant isn't attested.

(417) **wetl?epi:teki, **wetl?epi:taki 'root' 094: (a) PA *weĉye:pitki, W wɨlápitk-, (b) Y ?wɨ?łpitɨk. The initial **we- marks the part-of-whole relationship and the its glottalization is analogical. The stem initial **e is subject to initial change in two of the branches: with archaic ablaut in Wiyot, and **-eg- 'iterative infix' 025 or 'plural' 338 in Algonquian (giving PA *y). It is lengthened as a result of the simplification of Pre-PA **tl (see Proulx 1984:193).

(418) **?ekweytp-, **?ekwayĉp- 'fear' 097: PA *kweĉpan-, Y ?ekweyłpel- 'be frightened'. PA *-an is an AI final, Y -el a passive one. Y we relfects **wa (with grade variation, see Proulx 1984:sec.4.2). This reconstruction, somewhat simpler than the original, suggests that **yt may be a source for PA *ĉ in clusters.

(419) **-_etekwl_- verb medial 'think': PA *-_te:h_ 'think' (F _iŝite:he:wa_ 'she thinks thus', Ps _alitahasu_ 'she thinks about her own welfare', Mc _et'lite:t'm_ 'she believes it' and _tel'ta:sit_ 'she thinks so'), W _natɨtw_- 'think highly of' beside _nat_- 'be big' (Reichard 1925:127, in normalized orthography).

## Appendix B: THE RITWAN HYPOTHESIS

Although Proto-Algic is composed of three main branches (Algonquian, Wiyot, and Yurok), originally in 1913 only the last two were recognized as being related and their protolanguage was called Proto-Ritwan (Teeter 1964:v). Their genetic relationship to Algonquian was recognized later the same year, and the protolanguage shared by the three was called Proto-Algonkin and later Proto-Algic. For a long time, evidence was far too scanty to determine whether Proto-Ritwan was the same language as Proto-Algic (i.e., whether the three Algic languages are related at the same time depth), or whether Ritwan was a subgrouping within Algic. The latter view came to be called 'the Ritwan Hypothesis'.

The Ritwan Hypothesis was originally associated with the idea that the Proto-Algic homeland was in the East (the Great Lakes area) and that the Wiyot and Yurok migrated to northwestern California (where they were territorially adjacent). Given an Eastern homeland, it does indeed seem unlikely that there would be two independent migrations to this corner of California. However, if the Algic homeland was in the West - especially in or near Northwestern California - there is nothing odd about the Wiyot and the Yurok independently ending up where they are.

Methodologically, one must assume that Algic is composed of three equal branches unless a Ritwan grouping can be proved - and no such proof exists. Berman (1982a:sec.3) did speak of a shared merger of Proto-Algic *_t_ and *_ĉ_ (which he wrote *_ꝁ_) and loss of vowel length as evidence suggesting a possible Wiyot-Yurok subgrouping. However, subsequent research involving many more cognate sets has clearly shown (Proulx 1984:175, 182) that *_ĉ_ and *_t_ are consonant-grade variants of each other - giving only the illusion of correspondences in some sets - and that the loss of vowel length follows several independent vowel mergers in the two languages and thus postdates their genetic separation.

Anticipating another argument, the slightly higher rate of basic vocabulary shared by California Algic (Proulx 1982a:191, 199) does not provide a basis for genetic grouping. Wiyot and Yurok are in contact, and languages in contact retain more shared vocabulary (Swadesh 1971:32).

Since there is no evidence for a Wiyot-Yurok subgrouping, Wiyot/Yurok cognate sets require a Proto-Algic etymology (unless borrowing appears likely).

## COMPREHENSIVE PROTO-ALGIC INDEX

The following index is a cumulative one for all the papers in this series. Reconstructions have been updated to reflect doublets, as well as the improvements in the phonology over the years (see the discussion of **i in the present paper, and Proulx [1992:11-19] for a summary of the other changes). We may conveniently refer to items in this index by their initial number with "i" (for "index") prefixed, e.g., i001, or alternatively by their old numbers (e.g., 208).

1. **-?eteyeke 'hortative' 208
2. **-?i:na:me 'imperative with first person object' 206
3. **-?o:nĉh 'by boat' 151
4. **-?seni, **-?sa:ni, **-V?s 'hand' 164
5. **-a:m? 'by smell' 139
6. **-a:wegec 'sound' 171
7. **-ag, **-eg 'subordinative theme' 334
8. **-ałw 'plant' 309
9. **-aw 'benefactive' 188
10. **-awê?m? 'companion or relative' 170
11. **-ay 'skin ' 053
12. **-ayeł 'liquid' 310
13. **-ayewi, **-eyewi 'spacial adlocative (toward)' 337

BEST COPY AVAILABLE

14. **-ckit-, **-ĉkit- 'throat' 152
15. **-ê 'adverbial locative' 268
16. **-eCew 'reflexive' 192, 386b
17. **-eg- 'iterative infix' 025, 'plural' 338
18. **-êhk 'do, make' 136, cf.  **-Vˆhk 'do, make' 157
19. **-ehtl, **-ehĉr 'by ear' 142
20. **-eka:t, **-tka:ti, **-ĉka:ĉi 'leg' 143
21. **-ekanĉeg, **-tkanĉega, **-ĉkanĉega 'fingernail' 134, 141
22. **-em 'relational' 179
23. **-em 'subordinative theme' 335
24. **-ephe:?w 'rain' 51, 144
25. **-etekwl 'think' 419
26. **-eThey, **-aT̄hay 'belly' 095, 145, 363
27. **-ew 'middle reflexive' 172
28. **-hpa:le:we 'front part' 312
29. **-i:ɬkw 'sleep, dream' 146
30. **-i:s 'paddle' 180
31. **-i:wel, **-i:war 'load, burden' 147
32. **-k, **-ɬk 'chest (body part)' 140, 343
33. **-Ka 'postclitic' 240-241, 243, 404
34. **-Kila(wa) 'personal pronoun' 364
35. **-li:ni 'eye' 045
36. **-ohl?, **-ohr? 'on foot' 148
37. **-o:?l? 'in a body of water' 149
38. **-o:l 'swim' 173

39. **-o:tetkohCriw-, **-o:tetkohTliw- 'kidney' 314

40. **-seni 'breast' 002

41. **-skw-Vtkani 'neck' 267, cf. 306, 165

42. **-so:Kani 'hip', **-ĉo:Kani 'buttocks' 133, 387

43. **-tapti:take, **-tapti:teke 'backbone' 326

44. **-tecake, **-tegeteke 'buttocks' 327

45. **-tekwli, **cekwri 'heart' 112

46. **-tka:ti, **-ĉka:ĉi, **-eka:t 'leg' 143

47. **-tkan 'abstract final' 306

48. **-tkanĉega, **-ĉkanĉega, **-ekanĉeg 'fingernail' 134, 141

49. **-tko?wi, **-cko?wi 'blood' 126

50. **-tlo: or **-tlew 'mouth' 031

51. **-tpeyekhani 'rib' 091

52. **-T 'person, third...' Goddard (1991:65); 'nominalizer' 299
< 'the one WH-' 253, 259

53. **-Vˆ 'nominalizer' 301

54. **-Vĉk, **-tk 'augmentative-pejorative' 291

55. **-V? 'tree, stem' 308

56. **-V?ep 'by tooth, bite, eat' 158

57. **-Vp 'liquid' 160

58. **-V?s 'hand' 164, 'hand, lower arm or leg' 016

59. **-VTw or **-VTew 'by tool, instrument, medium' 187, 369

60. **-eTew 'reciprocal' 193, 368a

61. **-Vc 'come, go' 174

62. **-Vckit, **-Vĉkit 'throat' 152

63. **-Vˆhk, **âhk-, 'do, make' 157, cf. **-êhk 'do, make' 136
64. **-Vhl, **-Vhr 'lie, fall' 153, 340
65. **-Vhli:m 'to bed' 153
66. **-Vk 'postradical extension' 354
67. **-Vkamy 'natural body of water' 154
68. **-Vkani 'instrument or product', cf. 292
69. **-VKa:p 'stand' 381
70. **-VKh 'long thing' 155, 'tree, stem' 308; 405
71. **-Vkws 'wind' 156
72. **-VKw 'person, second...plural' 398
73. **-VK 'person, third ... subject' 293, 300, 396b; 'nominalizer' 292, 300, 403
74. **-Vl 'fly' 175
75. **-Vl 'mediopassive' 176
76. **-Vl, **-Vr 'nominalizer' 302
77. **-Vl, **-Vr 'break, cut' 183
78. **-Vlt, **-Vrt '1-dimensional' 281
79. **-Vɬ 'person, third ... subject' 201; 'nominalizer' 298
80. **-Vm 'by feeling, thought, speech, mouth, eating' 181
81. **-Vm 'nominalizer' 294
82. **-Vm 'postradical extension' 355
83. **-Vm 'relational' 191
84. **-Vm 'person, third-... subject' 295
85. **-Vn 'exerting fine control with the hand' 182
86. **-Vn 'nominalizer' 296
87. **-Vn 'nondeliberate action' 177

88. **-Vn 'postradical extension' 356

89. **-Vn or **-Vnag 'indefinite subject' 297

90. **-Vˆnc 'diminutive' 137, 189, 290

91. **-Vnki 'punctual spacial locative (in, at)' 336

92. **-Vp-eT 'put something in place' 161, 406

93. **-Vpiti 'tooth, teeth' 076, 159

94. **-VpiT 'tie, string, root' 162, 412 **apiT-, **epiT- 'tie, cord' 347, 410

95. **-Vpl, **-Vpr 'hair' 163

96. **-Vs 'causitive' 184

97. **-Vs, **-Vŝ 'second person object' 196

98. **-Vs-aKoK, **-Vŝ-eKoK 'I-you pl.' 384

99. **-Vskw 'neck' 165, cf. 267

100. **-Vŝ 'head' 166

101. **-Vt 'foot' 167

102. **-VT 'second person subject' 212, 392

103. **-VT 'nonpersonal object' 178, 406

104. **-Vt, **-Vc 'by heat' 185

105. **-Vtk, **-Včk 'body, leg, foot' 168

106. **-VTl 'sinister, false' 169

107. **-VTl-VKha, **-VCr-VKha, **-VCˆr-VKha 'digit from five to ten' 109, 415

108. **-VT 'second person subject' 212, 397

109. **-VTw 'by tool, instrument, medium' 187, 369

110. **-Vˆw 'be' 138, 190

111. **-Vw 'nominalizer' 304

112. **-Vw 'postradical extension' 357

113. **-Vw 'reflexive' 192

114. **-Vw '[abstract, often used with instrumentals and body-part medials]' 186

115. **-Vy 'indefinite subject' 305

116. **-Vy 'mediopassive' 194

117. **-Vy 'postradical extension' 358

118. **-Vyaki 'kind, sort, way, place, time' 311

119. **-wan, **-wen 'abstract final' 307

120. **-wi:wa 'wife', **wi:wali 'his wife' 313

121. **Øa:-, **Øakw- 'go' 220

122. **?e:neca 'shell, bivalve ...' 084

123. **?eki?yem- 'hat' 047

124. **?ekw- 'same, just exactly the time, manner, thing' 379

125. **?ekweytp-, **?ekwayĉp- 'fear' 097, 418

126. **?entlekwi, **?enetlekw- 'wonder, I ... if' 119

127. **a:?lewe 'projectile point' 315b

128. **a:l-, **e:r-, **-e:l- 'think, feel' 254

129. **a:p-, **era:p- 'duplicate' 012

130. **a:pł- 'sinew' 104

131. **a:wal- 'around' 056

132. **aTl- 'tell experiences' 116

133. **aTlemliyo:lkw- 'tell an origin myth' 115

134. **âhk-, **-Vˆhk 'do, make' 136, 157

135. **ahp-, **ehp-, **ahpi:g-, **ehpi:g- 'press down on, bruise' 339

136. **ehT- **ahT-, **ahC(am)-, **ahC^(am)- 'give (especially food)' 011, 370

137. **aleTewk-, **a:laTewk-, **a:leTewk-, 'be a cloud or shadow' 395

138. **apiT-, **epiT- 'tie, cord' 347, 410 **-VpiT 'tie, string, root' 162, 412

139. **apiTe:kh-, **-a:piTe:kh 'cord' 348, 411

140. **apiTeTk(w)-, **epiTeTk(w)- 'tie into a round shape, i.e., a bunch or bundle' 346, 389

141. **as?ola:?w- 'pound' 120

142. **asewe 'projectile point' 315a

143. **aSlek-, **aSrak-, **-erek-, **-arek- 'skin' 280

144. **ata:gw, **eta:gw 'belly' 283

145. **atkehk?wa 'pot' 028

146. **atk-, **atkeyi, **aĉkeyi, **-atk 'earth, land' 349

147. **cawoni 'fishnet' 341

148. **ĉep?eɬ-, **ĉa:p?eɬi, **-ĉp?eɬe 'eyelash' 325

149. **ĉo:la 'aunt, maternal ...' 048

150. **e?m?i:Ka 'pigeon' 121, 386

151. **ek?ey-, **ek?ay- 'conceal, hide' 122

152. **el-, **al- 'be such a number over four' 256

153. **elka:yome, **erka:yome 'bat (animal)' 320

154. **er-, **ar-, **el- **al- 'thither, thus, that way, like that, that sort' 255

155. **era:p-, **a:p- 'duplicate' 012

156. **i:m- 'hurriedly' 123

157. **i:?meli 'below' 124

158. **ihkwa 'louse' 009

159. **kataw 'want to, be inclined to' 054

160. **K(e)-, **Ket-, **Ke?- 'person, second...' 129, 366

161. **kec-, **ka:c- 'wipe' 350

162. **keĉhay- 'day, be ...' 017

163. **kega:rKa, **kaga:lKa 'gull' 135, 385b

164. **KeKila(wa) 'second person singular personal pronoun' 385b

165. **KeKilawa:wa, **KeKilewa:wa 'ye' 132, 261, 364

166. **kela:hkewa, **kegela:hkewa 'goose, geese' 106

167. **kelom- 'turn' 049

168. **kełeyet- 'hot' 093

169. **kemotl- 'steal' 021

170. **ken-, **kan- 'motionless, left ...' 080

171. **kenew- 'long' 077

172. **kenlegewa 'hawk species brownish' 075

173. **kep-e?- 'close an opening' 407

174. **Kep-etkoy?- 'choke on something (food, tobacco smoke)' 408

175. **kep-w- 'cover' 409

176. **keSt-, **keSĉ- 'big, much' 272

177. **ketp-, **keĉp- 'thick cloth, etc.' 040

178. **kewo?m- 'back 180 degrees' 351

179. **keyew?a:tl- 'turn treacherously' 225

180. **kha, **khe 'future of obligation' 043, 248

181. **ki:m-, **kegi:m- 'bad, illicit' 061

182. **ki:mla:tl- 'bad, be ...' 107

183. **ki:t-, **ki:ĉ-, **ki:ĉi- 'finish, complete' 007

184. **ki:tep-, **ki:tap-, **ki:cep- 'finish eating' 100

185. **ki 'previous action' 250

186. **ki, **khi 'can, able to' 249

187. **Kiyew- 'in a circle, around, about' 382

188. **ko:lew-, **ko:law- 'add to, mix in' 226

189. **kwa, **kwe 'foreground' 252

190. **kwetk- 'angle (for fish)' 342

191. **kweyeC^kweyehr-, **kweyeC^kweyehl- 'whistle' 088 393

192. **k- 'the one previously mentioned' 251

193. **k-, **ka 'nominal negator, prohibitive' 262

194. **le:k?awi 'sand', **lege:k?ewi 'sandbar' 072

195. **leski:m-, **leŝki:m-, **la:ŝki:m- 'scold' 041

196. **leweTk- 'soft' 085 391

197. **lewk-, **legewk- 'overcast or covered, be ...' 149

198. **lo:yewa:s- 'blow or suck with a pipe' 070

199. **lo:yewekws- 'blow, the wind ...-s' 071; **-Vkws 'wind' 156

200. **m- 'indefinite possessor' 131

201. **m- **nVm- **ma 'general verbal negator' 264

202. **ma:tkwe, **-Vĉkw 'head' 329

203. **ma 'proximal time' 244

204. **ma, **me 'go and do' 258

205. **megepl-, **megepr- 108, **-Vpl, **-Vpr 'hair' 163

206. **megwi:Stl-, **megwi:Scr- 'bare, without anything' 273

207. **mehŝe 'fire' 332

BEST COPY AVAILABLE

208. **mekwehca, **megekwehca 'snail' 033
209. **mekw- 'bark, chatter' 105
210. **meleK-, **mer- 'rub hard' 344
211. **men-, **menehkw-, **menahkw- 'drink' 228
212. **mena, **megena 'berry, berries' 073
213. **merkwe, **markwe 'grass, herb, medicine' 330
214. **metlag-, **megetleg-, **megecreg- 'dung' 052
215. **mew-, **megew- 'eat' 139
216. **mew?om- 'come from there' 057
217. **mey?eɬki 'nettle' 062
218. **meyehkhwel-, **meyehkhwali 'weep, mourn; tear' 089
219. **meɬi 'emphatic' 023
220. **mikhw- 'salalberry' 032
221. **na:pe:gwa 'male' 284
222. **na:T- 'large' 374
223. **na:wal?- 'following one behind the other' 055
224. **na 'distal time' 245
225. **naKh- 'be skillful' 229, 388
226. **nawe:l- 'good-looking' 060
227. **ne:Sc- 'weak, inadequate' 274
228. **ne:St- 'exhausted, to exhaustion' 274
229. **ne:w- 'see' 042
230. **neg-, **nag- 'carry' 059
231. **negi:m- 'carry away food in mouth' 059
232. **negom?- 'carry on back' 059

233. **nehk-, **negehk- 'call, name' 127

234. **neKila(wa) 'first person singular personal pronoun' 385a

235. **nekwet-, **nekwec- 'one' 063

236. **nen?-, **negen?- 'see, look for', bipersonal **nen?aw-, **negen?aw- 230

237. **nepe:gw- 'be male' 284

238. **nepek-, **negepek- 'be bloody, red' 125

239. **nepi?yi 'water' 067

240. **neSĉ-, **neSt- 'night; dark, lonely, awful' 276

241. **neSĉV 'last night' 275

242. **n(e)-, **net-, **ne?- 'person, first...' 128, 365

243. **newon-, **newona:ĉ- 'suckle' 006

244. **ni 'distal space' 256

245. **nikhl-, **nikhr- 'three' 046

246. **niTema:K- 'two days' 065, 373

247. **niTeThey-, **niTeC^hey- 'two humans' 036 372

248. **niye:?w- 'four' 050

249. **no:K- 'such a distance' 383

250. **no:Tpen- 'pursue' 098, 371

251. **pe?meyi 'grease' 024

252. **pegemi 'knife', **pegemi:pi 'knife-making material' 114

253. **epehtl- or **epentl- 'listen' 018

254. **pekw-, **pakw- 'into a body of water' 233

255. **pekw-, **pekwen-, **pakwen- 'pierce through an obstacle' 232

256. **pekwan- 'rhubarb, indian ...' 069

257. **pel- 'big' 234

258. **pelakw-, **pelkw- 'peel' 231

259. **pele:gwa 'large bird' 285

260. **pełełkewe, **-Vpełełkewe 'rock' 316

261. **penekwi 'powder' 083

262. **peni:wal-, **peni:war- 'drop a back-pack' 082

263. **peTk-, **peCk- 'catch fire' 376

264. **pew- 'put on the fire' 140, 352

265. **phele 'flint' 323

266. **pipo:ne 'winter' 321

267. **pitekwli, **pitekwali, **piĉekwri 'basket' 111

268. **ra, **la 'post-negative enclitic' 265

269. **re?w-, -e?w 'catch, kill' 345

270. **rente or **sente, **ŝente 'evergreen tree, brush, or bough' 333

271. **rey(?)-, **ley(?)- 'ropelike' 319

272. **s?e:gw... 'madrone' 286

273. **s?yoyeTkw-, **s?yoyeCkw-, **s?yoyeC^kw- 'slip, be smooth' 086 392

274. **sa:p-, **ŝa:p-, or **ra:p- 'repeat again' 012

275. **sema:lt-, **ŝema:lt- 'bow' 282

276. **SleSl-, **SlaSl-, **SleSr-, **SraSl-, **SlegeSl- 'divide' 277

277. **Sleyep?et-, **Sleyep?at-, **Sreyep?ec- 'singe' 278

278. **ŝegw-, **ŝagw-, **segegw- 'be tired of' 287

279. **t-, **ta, **ti 'the one known but not previously mentioned' 253

BEST COPY AVAILABLE

280. **T- 'third person subject' 293, 300, 396a

281. **T- 'be, exist, dwell' 221-222, 399 < 'the one WH-' 253, 259

282. **T- 'immobilizer: in one place' 400

283. **T- 'WH-, TH-' 402; **Tâ:ł?- **Tâ:ł- 'WH-?' 263

284. **Ta:- 'dwell' 222

285. **Ta:wal- 'be around' 377

286. **Takwi 'there is; she or it exists, is located, or dwells there' 200

287. **Tar-, **Tel- 'there' 223, 401

288. **Tehkon- 'grasp, hold onto something with the hand' 064, 019, 360

289. **tekw-, **takw- 'pound' 014

290. **tekw-, **takw- 'together, grouped' 380; **?ekw- 'same, just exactly the time, manner, thing' 379

291. **tek-, **tak- 'few' 235

292. **tem- 'cut or break in lengths or leaving a stubble or stump' 081

293. **tep-, **tegep- 'stand, fix upright' 236

294. **tepehtl- 'listen' 092

295. **tepehtleKa 'ear' 092, 403

296. **tetkw-, **tatkw-, **tačkw- 'short, be ...' 079 390

297. **teyew?- 'burn something' 117

298. **thigw- 'through an opening; out' 288

299. **thel-, **thar- 'talk' 015, 237

300. **thi?ema:?i 'elderberry bush' 066

301. **ti:kw- 'break' 087

302. **Tik- 'urinate' 008, 099, 359

303. **wa:kel- 'peppernuts' 101

304. **wa:wa?lewi 'egg' 003

305. **wa:yelKwa or **?wa:yelkwa 'roe' 414

306. **wa 'this personal, extended' 240

307. **wałanyi 'tail' 022

308. **wanri?a:gwi, **wenli?a:gwi 'coals, charcoal' 289

309. **wareyi, **wegeleyi 'navel' 110

310. **waSlak-, **waSlek-, **waSrak-, **-erek- and **-arek- 'skin' 280

311. **way- 'overnight' 353

312. **we 'this nonpersonal, extended' 241

313. **we, **wi 'emphatic' 263

314. **wechowe, **-thow 'quadruped-tail' 317

315. **weĉekwani, **weĉegekwani 'fish-tail' 074

316. **weĉekwane, **wegeĉekwane, **wecekwane 'elbow' 322

317. **wehTelkweni, **wehTelkwani, **wehTerkwani 'branch' 038, 362

318. **weKila(wa) 'third person singular personal pronoun' 385c

319. **wêkwi 'what?' 264

320. **wel-, **welakhw-, **wegel-, **wegere 'fat' 324

321. **welekwe:ŝkwi, **walekwe:ŝkwi, **walakwe:skwi 'tree bark' 113

322. **wełkani 'bone' 026

323. **wełkweni, **wełkwani 'liver' 035

324. **wen-, **wegen-, **negen- 'name, mention by ...' 058

325. **weskiTyi 'pipestem, hollow tube used in smoking' 331, 361

326. **weSlaw-, **wa:Slaw- 'far away' 279

327. **w(e)-, **wet-, **we?- 'person, third...' 130, 367

328. **wetag-, **negeĉag- 'behind' 238

329. **weTa(:)phega 'root, fine ... used in sewing (usually spruce)' 102, 364

330. **weTayehca 'dog, (her)...' 413

331. **weTehpehKa 'backbone, (her)...' 378

332. **wetempi 'head' 090

333. **weTewi, **wegeTawi 'flesh, body, (her)...' 375

334. **we?ilkw-, **we?irkw- OR **w?erkw-, **w?egerkw-, **w?egelkw- 'bundle-strap' 037, 416

335. **wetḱanĉega, **weĉkenĉega 'hoof, claw' 039

336. **wetl?epi:teki, **wetl?epi:taki 'root' 094, 417

337. **wi 'future of volition' 044, 247

338. **wo 'this restricted' 239

339. **wVˆł?a 'who?' 265

340. **ya 'that personal, extended, then' 243

341. **yeTkw-, **yaTkw- 'curve' 394

342. **yeTekwe 'maggot' 318

343. **yo 'that restricted' 242

344. **y?o:nĉhi, **y?o:neĉh- 'boat' 118, medial **-?o:nĉh 151

# INDEX OF GLOSSES

1-dimensional. 281 **-Vlt, **-Vrt

abstract final. 306 **-tkan

abstract final. 307 **-wan, **-wen

abstract, often used with instrumentals and body-part medials. 186 **-Vw

add to, mix in. 226 **ko:lew-, **ko:law-

adverbial locative. 268 **-ê

angle (for fish). 342 **kwetk-

around. 056 **a:wal-

augmentative-pejorative. 291 **-Vĉk, **-tk

aunt, maternal .... 048 **ĉo:la

back 180 degrees. 351 **kewo?m-

backbone, (her).... 378 **weTehpehKa

backbone. 326 **-tapti:take, **-tapti:teke

bad, be .... 107 **ki:mla:tl-

bad, illicit. 061 **ki:m-, **kegi:m-

bare, without anything. 273 **megwi:Stl-, **megwi:Scr-

bark, chatter. 105 **mekw-

basket. 111 **pitekwli, **pitekwali, **piĉekwri

bat (animal). 320 **elka:yome, **erka:yome

be around. 377 **Ta:wal-be male. 284 **nepe:gw-

be skillful. 229, 388 **naKh-

be such a number over four. 256 **el-, **al-

be tired of. 287 **ŝegw-, **ŝagw-, **segegw-

be, exist, dwell. 221-222, 399 **T-

be. 138, 190 **-Vˆw

behind. 238 **wetag-, **negeĉag-

belly. 095, 145, 363 **-eThey, **-aThay

belly. 283 **ata:gw, **eta:gw

below. 124 **i:?meli

benefactive. 188 **-aw

berry, berries. 073 **mena, **megena

big, much. 272 **keSt-, **keSĉ-

big. 234 **pel-

blood. 126 **-tko?wi, **-cko?wi

blow or suck with a pipe. 070 **lo:yewa:s-

blow, the wind ...-s. 071 **lo:yewekws-

boat. 118, medial **-?o:nĉh 151 **y?o:nĉhi, **y?o:neĉh-

body, leg, foot. 168 **-Vtk, **-Vĉk

body, flesh, (her)... 375 **weTewi, **wegeTawi

bone. 026 **weɫkani

bow. 282 **sema:lt-, **ŝema:lt-

branch. 038, 362 **wehTelkweni, **wehTelkwani, **wehTerkwani

break, cut. 183 **-Vl, **-Vr

break. 087 **ti:kw-

breast. 002 **-seni

bundle-strap. 037, 416 **we?ilkw-, **we?irkw- OR

bundle-strap. 037, 413 **w?egelkw-, **w?erkw-, **w?egerkw-

burn something. 117 **teyew?-

buttocks. 133 **-ŝo:kani

buttocks. 327 **-tecake, **-tegeteke

by boat. 151 **-?o:nĉh

by ear. 142 **-ehtl, **-ehĉr

by feeling, thought, speech, mouth, eating. 181 **-Vm

by heat. 185 **-Vt, **-Vc

by mouth. 031, 150 **-o:ł, **-tlo:łi

by smell. 139 **-a:m?

by tool, instrument, medium. 187, 369 **-VTw or **-VTew

by tooth, bite, eat. 158 **-V?ep

call, name. 127 **nehk-, **negehk-

can, able to. 249 **ki, **khi

carry away food in mouth. 059 **negi:m-

carry on back. 059 **negom?-

carry. 059 **neg-, **nag-

catch, kill. 345 **re?w-, -e?w

catch fire. 381 **peTk-, **peCk-

causitive. 184 **-Vs

chest (body part). 140, 343 **-k, **-łk

choke on something (food, tobacco smoke). 408 **Kep-etkoy?-

circle, in a...; around, about. 382 **Kiyew-

close an opening.  407 **kep-e?-

cloud or shadow.  395 **aleTewk-, **a:laTewk-, **a:leTewk-

coals, charcoal.  289 **wanri?a:gwi, **wenli?a:gwi

come from there.  057 **mew?om-

come, go.  174 **-Vc

companion or relative.  170 **-awê?m?

conceal.  122 **ek?ey-, **ek?ay-

cord.  348, 411 **ap(-iT)-e:Kh-, **-a:p(-iT)-e:Kh

cover.  409 **kep-w-

curve.  394 **yeTkw-, **yaTkw-

cut or break in lengths or leaving a stubble or stump.  081 **tem-

day, be ....  017 **keĉha:y-

digit from five to ten.  109, 415 **-VTl-VKha, **-VCr-VKha, **-VC^r-VKha

diminutive.  137, 189, 290 **-V^nc

distal space.  256 **ni

distal time.  245 **na

distance, such a...  383 **no:K-

divide.  277 **SleSl-, **SlaSl-, **SleSr-, **SraSl-, **SlegeSl-

do, make.  136 **-êhk

do, make.  136, 157 **âhk-, **-V^hk

dog, (her)....  413 **wetayehca

drink.  228 **men-, **menehkw-, **menahkw-

drop a back-pack.  082 **peni:wal-, **peni:war-

dung.  052 **metlag-, **megetleg-, **megecreg-

dwell.  222 **Ta:-

ear. 092, 403 **tepehtleKi or **tepentleKi

earth, land. 349 **atk-, **atkeyi, **aĉkeyi, **-atk

eat. 139 **mew-, **megew-

egg. 003 **wa:wa?lewi

elbow. 322 **weĉekwane, **wegeĉekwane, **wecekwane

elderberry bush. 066 **thi?ema:?i

emphatic. 023 **meƚi

emphatic. 263 **we, **wi

evergreen tree, brush, or bough. 333 **rente or **sente, **ŝente

exerting fine control with the hand. 182 **-Vn

exhausted, to exhaustion. 274 **ne:St-

eye. 045 **-li:ni

eyelash. 325 **ĉep?eƚ-, **ĉa:p?eƚi, **-ĉp?eƚe

far away. 279 **weSlaw-, **wa:Slaw-

fat. 324 **wel-, **welakhw-, **wegel-, **wegere

fear. 097, 418 **?ekweytp-, **?ekwayĉp-

few. 235 **tek-, **tak-

fingernail. 134, 141 **-tkanĉega, **-ĉkanĉega, **-ekanĉeg

finish eating. 100 **ki:tep-, **ki:tap-, **ki:cep-

finish, complete. 007 **ki:t-, **ki:ĉ-, **ki:ĉi-

fire. 332 **mehŝe

first person singular personal pronoun. 385a **neKila(wa)

fish-tail. 074 **weĉekwani, **weĉegekwani

fishnet. 341 **cawoni

flesh, body, (her)... **weTewi, **wegeTawi flint. 323 **phele

fly. 175 **-Vl

following one behind the other. 055 **na:wal?-

foot. 167 **-Vt

foreground. 252 **kwa, **kwe

four. 050 **niye:?w-

front part. 312 **-hpa:le:we

future of obligation. 248 **kha, **khe

future of volition. 247 **wi

general verbal negator. 264 **m- **nVm- **ma

give (especially food). 11, 370 **ehT- **ahT-, **ahC(am)-, **ahCˆ(am)-

go and do. 258 **ma, **me

go. 220 **Øa:-, **Øakw-

good-looking. 060 **nawe:l-

goose, geese. 106 **kela:hkewa, **kegela:hkewa

grasp, hold onto something with the hand. 064, 019, 360 **Tehkon-

grass, herb, medicine. 330 **merkwe, **markwe

grease. 024 **pe?meyi

gull. 135, 385b **kega:rKa, **kaga:lKa

hair. 163 **-Vpl, **-Vpr

hair. 163 **megepl-, **megepr- 108, **-Vpl, **-Vpr

hand. 016, 164 **-V?s, **-?seni, **-?sa:ni 'hand, lower arm or leg'

hat. 047 **?eki?yemhawk species brownish. 075 **kenlegewa

head. 090 **wetempi

head. 166 **-Vŝ

head. 329 **ma:tkwe, **-Vĉkw

heart. 112 **-tekwli, **cekwri

hip. 133, 387 **-so:Kani

hoof, claw. 039 **wetkanĉega, **wečkenĉega

hortative. 208 **-?eteyeke

hot. 093 **kełeyet-

hurriedly. 123 **i:m-

I-you pl. 384 **-Vs-aKoK, **-Vŝ-eKoK

immobilizer: in one place. 400 **T-

imperative with first person object. 206 **-?i:na:me

in a body of water. 149 **-o:?l?

indefinite subject. 297 **-Vn or **-Vnag

indefinite subject. 305 **-Vy

instrument or product', cf. 292 **-Vkani

into a body of water. 233 **pekw-, **pakw-

iterative infix. 025, 'plural' 338 **-eg-

kidney. 314 **-o:tetkohCriw-, **-o:tetkohTliw-

kind, sort, way, place, time. 311 **-Vyaki

knife. 114 **pegemi

knife-making material. 114 **pegemi:pi

large bird. 285 **pele:gwa

large. 374 **na:T-

last night. 275 **neSĉv

leg. 143 **-tka:ti, **-ĉka:ĉi, **-eka:t

lie, fall. 153, 340 **-Vhl, **-Vhr

liquid. 160 **-Vp

liquid. 310 **-ayeł

listen. 018, 092 **epehtl-, **tepehtl-, **era:pehtl-

liver. 035 **wełkweni, **wełkwani

load, burden. 147 **-i:wel, **-i:war

long thing. 155, 405 **-VKh

long. 077 **kenew-

louse. 009 **ihkwa

madrone. 286 **s?e:gw...

maggot. 318 **yeTekwe

male. 284 **na:pe:gwa

mediopassive. 176 **-Vl

mediopassive. 194 **-Vy

middle reflexive. 172 **-ew

motionless, left .... 080 **ken-, **kan-

mouth. 150 **-tlo:, or **-tlew 31, **-o:ł 150

name, mention by .... 058 **wen-, **wegen-, **negen-

natural body of water. 154 **-Vkamy

navel. 110 **wareyi, **wegeleyi

neck. 165, cf. 267 **-Vskw

neck. 267, cf. 306, 165 **-skw-vtkani

nettle. 062 **mey?ełki

night; dark, lonely, awful.  276 **neSĉ-, **neSt-

nominal negator, prohibitive.  262 **k-, **ka

nominalizer.  292, 300, 403 **-VK

nominalizer.  294 **-Vm

nominalizer.  296 **-Vn

nominalizer.  298 **-Vł

nominalizer.  299 **-t

nominalizer.  301 **-Vˆ

nominalizer.  302 **-Vl, **-Vr

nominalizer.  304 **-Vw

nondeliberate action.  177 **-Vn

nonpersonal object.  178, 406 **-VT

on foot.  148 **-ohl?, **-ohr?

one.  063 **nekwet-, **nekwec-

overcast or covered, be ....  149 **lewk-, **legewk-

overnight.  353 **way-

paddle.  180 **-i:s

peel.  231 **pelakw-, **pelkw-

peppernuts.  101 **wa:kel-

person, first....  128, 365 **n(e)-, **net-, **ne?-

person, second...plural.  398 **-VKw

person, second...subject.  212, 397 **-VT

person, second....  129, 366 **K(e)-, **Ket-, **Ke?-

person, second....singular personal pronoun.  385b **KeKila(wa)

person, third ... subject.  293, 300, 396b **-VK

person, third ... subject. 201 **-Vł

person, third-... subject. 295 **-Vm

person, third.... subject. 396a **-T

person, third.... 130, 367 **w(e)-, **wet-, **we?-

personal pronoun. 364 **-Kila(wa)

pigeon. 121, 386 **e?m?i:Ka

pierce through an obstacle. 232 **pekw-, **pekwen-, **pakwen-

pipestem, hollow tube used in smoking. 331, 361 **weskiTyi

plant. 309 **-ałw

post-negative enclitic. 265 **ra, **la

postclitic. 240-241, 243, 404 **-Ka

postradical extension. 354 **-Vk

postradical extension. 355 **-Vm

postradical extension. 356 **-Vn

postradical extension. 357 **-Vw

postradical extension. 358 **-Vy

pot. 028 **atkehk?wa

pound. 014 **tekw-, **takw-

pound. 120 **as?ola:?w-

powder. 083 **penekwi

press down on, bruise. 339 **ahp-, **ehp-, **ahpi:g-, **ehpi:g-

previous action. 250 **ki

projectile point. 315b **a:?lewe

projectile point. 315a **asewe

proximal time. 244 **ma

punctual spacial locative (in, at).  336 **-Vnki
pursue.  098, 371 **no:Tpen-
put something in place.  161, 406 **-Vp-eT
put on the fire.  140, 352 **pew-
quadruped-tail.  317 **wechowe, **-thow
quote.  078 **ŝepeyer?-, **ŝegepeyer?-
rain.  144 **-ephe:?w
reciprocal.  193, 368a **-eTew
red.  125 **nepek-
reflexive.  192, 368b **-eCew
relational.  179 **-em
relational.  191 **-Vm
repeat again.  012 **sa:p-, **ŝa:p-, or **ra:p-
rhubarb, indian ....  069 **pekwan-
rib.  091 **-tpeyekhani
rock.  316 **pełełkewe, **-Vpełełkewe
roe.  414 **wa:yelKwa or **?wa:yelKwa
root, fine ... used in sewing (usually spruce).  102, 364 **weTa(:)phega
root.  094, 417 **wetl?epi:teki, **wetl?epi:taki
ropelike.  319 **rey(?)-, **ley(?)-
round shape, tie into a...  346, 389 **apiTeTk(w)-, **epiTeTk(w)-
rub hard.  344 **meleK-, **mer-
salalberry.  032 **mikhwałi
same, just exactly the time, manner, thing.  379 **?ekw-
sand.  072 **le:k?awi

sandbar. 072 **lege:k?ewi

scold. 041 **leski:m-, **leŝki:m-, **la:ŝki:m-

second person singular personal pronoun. 364b **KeKila(wa)

see, look for', bipersonal 230 **nen?aw-, **negen?aw-

see, look for. 230 **nen?-, **negen?-

see. 042 **ne:w-

shell, bivalve .... 084 **?e:neca

short, be .... 079 390 **tetkw-, **tatkw-, **taĉkw-

sinew. 104 **a:pł-

singe. 278 **Sleyep?et-, **Sleyep?at-, **Sreyep?ec-

sinister, false. 169 **-VTl

skin . 053 **-ay

skin. 280 **aSrak-, **aSlek-, **-arek-, **-erek-

skin. 280 **waSlak-, **waSlek-, **waSrak-, **-erek- and **-arek-

sleep, dream. 146 **-i:łkw

slip, be smooth. 086 392 **s?yoyeTkw-, **s?yoyeCkw-, **s?yoyeCˆkw-

snail. 033 **mekwehca, **megekwehca

soft. 085 391 **leweTk-

sound. 171 **-a:wegec

spacial adlocative (toward). 337 **-ayewi, **-eyewi

stand, fix upright. 236 **tep-, **tegep-

stand. 381 **-VKa:p

steal. 021 **kemotl-

subordinative theme. 334 **-ag, **-eg

subordinative theme. 335 **-em

suckle. 006 **newon-, **newona:ĉ-

swim. 173 **-o:l

tail. 022 **wałanyi

talk. 015, 237 **thel-, **thar-

tell an origin myth. 115 **aTlemliyo:lkw-

tell experiences. 116 **aTl-

that personal, extended, then. 243 **ya

that restricted. 242 **yo

the one known but not previously mentioned. 253 **t-, **ta, **ti

the one previously mentioned. 251 **k-

the one wh-. 253, 259 **T-

there is; she or it exists, is located, or dwells there. 200 **Takwi

there. 223, 401 **Tar-, **Tel-

thick cloth, etc.. 040 **ketp-, **keĉp-

think. 419 **-etekwl

think, feel. 254 **a:l-, **e:r-, **-e:l-

third person singular personal pronoun. 385c **weKila(wa)

this nonpersonal, extended. 241 **we

this personal, extended. 240 **wa

this restricted. 239 **wo

thither, thus, that way. 255 **er-, **ar-, **el- **al-

three. 046 **nikhl-, **nikhr-

throat. 152 **-Vckit, **-Vĉkit

through an opening; out. 288 **thigw-

tie into a round shape. 346, 389 **apiTetk(w)-, **epiTetk(w)-

tie, cord. 347, 410 **apiT-, **epiT-

tie, string, root. 162, 412 **-VpiT

tired of, be .... 287 **ŝegw-, **ŝagw-, **segegw-

to bed. 153 **-Vhli:m

together, grouped. 380 **tekw-, **takw-

tooth, teeth. 076, 159 **-Vpiti

tree bark. 113 **welekwe:ŝkwi, **walekwe:ŝkwi, **walakwe:skwi

tree, stem. 308 **-VKh

turn treacherously. 225 **keyew?a:tl-

turn. 049 **kelom-

two days. 065, 373 **niTema:K-

two humans. 036, 372 **niTeThey-

urinate. 008, 099, 359 **Tik-

want to, be inclined to. 054 **kataw

water. 067 **nepi?yi

weak, inadequate. 274 **ne:Sc-

weep, mourn; tear. 089 **meyehkhwel-, **meyehkhwali

WH-?. 263 **Tâ:ɬ?- or **Tâ:ɬ-

WH-, TH-. 402 **T-

what?. 264 **wêkwi

whistle. 088 393 **kweyeC^kweyehr-, **kweyeC^kweyehl-

who?. 265 **wV^ɬ?a

wife' 313 **-wi:wa

wife, his.... 313 **wi:wali

wind. 156 **-Vkws

winter. 321 **pipo:ne
wipe. 350 **kec-, **ka:c-
wonder, I ... if. 119 **?entlekwi, **?enetlekw-
ye. 261, 364 **KeKilawa:wa, **KeKilewa:wa

# THE LIMIT OF STRUCTURE PRESERVATION IN DAK$^h$OTA LEXICAL PHONOLOGY

John Kyle
University of Kansas

Abstract: Some of the earliest papers on Lexical Phonology claim that structure preservation applies throughout a Lexical derivation and may only be shut off by exiting the Lexicon. Work by Kellogg (1991) in Lak$^h$ota attempts to uphold this relationship between Lexical Phonology and Structure Preservation but recent work in Lexical Phonology and some older work in Dak$^h$ota refute this claim. After a minimal discussion of Dak$^h$ota phonology, morphology and how they relate to each other in Lexical Phonology, I will take up the problem of syllable structure within the Lexicon and show that Structure Preservation seems to be shut off early in the Lexical derivation.

## Introduction

The sources for this paper are the works on Dakota by Boas & Deloria (1941), Shaw (1980, 1985) and on Lak$^h$ota by Rood & Taylor (1976), Kellogg (1991). Although my sources come from both Lak$^h$ota and Dak$^h$ota, I will only use the term Dak$^h$ota unless referring to a specific text or rule. One reason for making this choice is that one of the most complete theoretical works was written by Pat Shaw (1980) on Dakota. She uses an *SPE* framework to develop Underlying Representations (UR) of many Dak$^h$ota words and since Lexical Phonolgy makes use of URs, it is only natural to use her book as a source.

Due to limited space and the large topic that I've chosen, I can't go into every aspect of the theories of Lexical Phonology, Prosodic Morphology/Phonology, or other Non-linear Phonological theories. Two excellent sources for more information are Autosegmental & Metrical Phonology by John A. Goldsmith (1990) and Morphological Theory by Andrew Spencer (1991).

Lak$^h$ota is a dialect of Dak$^h$ota (or Dakota), which is a member of the Upper-Mississippi River sub-family of the Siouan family. There are three dialects of Dak$^h$ota: Dak$^h$ota (d-dialect), Lak$^h$ota (l-dialect), and Nak$^h$ota (n-dialect). The Lak$^h$ota dialect is generally associated with the reservations west of the Missouri River in South Dakota: Pine Ridge, Rosebud, Lower Brule, etc. Dak$^h$ota is east of the Missouri in South Dakota and Minnesota. Nak$^h$ota is associated with the northern Sioux in Canada and North Dakota. Of the three dialects, Lak$^h$ota has the most speakers and since most of the major pedagogical texts are in Lak$^h$ota its use seems to be spreading.

*Kansas Working Papers in Linguistics*, Volume 19, Number 2, 1994, pp. 183-199

## Dak$^h$ota Phonemes

Table I is a compilation of the phonemes described by both Boas & Deloria (1941) and Shaw (1980). The (d) and (l) are shown in parenthesis since Dak$^h$ota uses /d/ where Lak$^h$ota uses /l/. When a /p/ or /k/ comes before an /m/ or /l/ it is realized phonetically as [b] or [g] respectively. Thus [b] is an allophone of /p/ and [g] is an allophone of /k/. Following Shaw (1980), I will not use them. [b] also occurs phonetically in first person singular forms of a group of verbs which have a stem initial *yu-* or *ya-* such as yuha 'to have.' The first person singular form would be (phonetically) [bluha] in Lak$^h$ota and [bduha] in Dak$^h$ota. There is a possibility that this occurence of [b] is the result of a sound change brought on by the first person pronominal prefix *wa-* and the initial [y] of the verb stem. In any case, the voiced stops usually occur in predictable environments so I will not treat them as phonemes. The [b] is shown only because Shaw (1980) mentions several words where it occurs outside the predictable environment.

| | labial | dental | palatal | velar | glottal | laryngeal |
|---|---|---|---|---|---|---|
| v-less | p | t | c | k | | |
| vless asp | $p^h$ | $t^h$ | $c^h$ | $k^h$ | | |
| v-less eject | p' | t' | c' | k' | | |
| voiced | (b) | (d) | | | | |
| vless fric | | s | š | x | | |
| eject fric | | s' | š' | x' | | |
| voiced fric | | z | ž | γ | | |
| nasals | m | n | | | | |
| glides | | (l) | y | w | ? | ħ |

**TABLE I**:(from Shaw (1980)) Dak$^h$ota Consonants

Table II shows the vowels for Dak$^h$ota, there are eight phonemic vowels in Dak$^h$ota: five oral and three nasal.

| | | |
|---|---|---|
| i,į | | u,ų |
| e | | o |
| | a,ą | |

TABLE II:(from Shaw (1980)) Dak$^h$ota Vowels

## Lexical Phonology

The theory of Lexical Phonology arose out of the SPE theory of phonology in an effort to explain problems with interaction of morphology and phonology. According to the SPE theory, morphological rules apply before phonological rules. In order to allow phonological rules to apply within a concatenated word, different classes of morphemes were given boundary markers (i.e. +,-,%,etc). A phonological rule could then be given a domain of application which would specify where the rule could apply. Lexical Phonology does away with the strict separation of morphology and phonology by allowing phonological rules to apply 'inbetween' morphological rules. The early papers on Lexical Phonology (Kiparsky, 1982, *inter alia.*) viewed it as a multi-level system in which a lexical item goes through derivations, inflections and sound changes. Each level consists of a morphological component followed by a phonological component thus enabling morphology and phonology to interact.

Each of the levels in this model is roughly equivalent in function to the boundaries used in SPE and each level is distinct from the other levels. The boundaries used in SPE are no longer needed since the phonological rules don't have to wait for all the morphological rules to apply. Also, processes from an earlier level and the morpheme boundaries it contained are not accessible to later levels. The convention used to ensure this inaccessibility is that of Bracket Erasure.

BRACKET ERASURE

Internal brackets are erased at the end of each level.

One problem that I've encountered in Lexical Phonology is the use of square brackets, [ ], to contain lexical material. Lexical representations are underlying or theoretical forms in contrast to phonetic representations which are surface forms. Yet phonetic forms are written with square brackets also. To avoid any confusion, all phonetic forms will specifically mentioned as such; any other use of square brackets will be for lexical material.

In the tradition of generative phonology, Shaw (1980) lists four types of boundaries for Dak$^{h}$ota:

| | | |
|---|---|---|
| Morpheme boundary | + | weak |
| Lexical derivational boundary | % | \| |
| enclitic boundary | = | \| |
| word boundary | # | strong |

Each of the first three boundaries is roughly equivelant in function to one of the levels in Shaw's (1986) Lexical Phonology model (see Table III).

Although the early versions of the theory viewed each level as being distinct and disallowed access to morphological information from an earlier level, the more recent versions have softened this stance. In the introduction to Kiparsky (1985), he goes so far as to refer to the levels as 'quasi-autonomous'. Mohanan (1986) refers to 'the loop' which permits the output from Level III to feed back into Level II. In English, this allows a compound (compounding is a Level III process) to acquire a Level II ending:

[half-hearted] → [[half-hearted][ness]].

Mohanan considers 'the loop' to be universal.

There are other aspects of Mohanan's version which make it stronger than Kiparsky's early version. Whereas Kiparsky can classify phonological rules as being lexical and post-lexical, Mohanan claims that it is the phonological rule's *application* which can be classified this way. Instead of two different sets of phonological rules, Mohanan has one set. Each rule is given a domain in which it applies. Rules may apply in the lexical module, the postlexical module, or in both. There are no 'lexical rules' but rather rules that apply within the lexicon. Rules are stated only once in the grammar but included are specifications regarding their relative ordering and domain of application. However, cyclicity is a property of the stratum *not* the rule. A rule may apply cyclically in a cyclic stratum and noncyclically in a noncyclic stratum. This contrasts with Kiparsky's claim of cyclic phonological rules.

### Dak$^{h}$ota Lexical Phonology

The first work published on Dak$^{h}$ota Lexical Phonology was by Shaw (1985) and I would direct the reader to it and the other articles in *Phonology Yearbook 2* which deal solely with Lexical Phonology (albeit slightly out of date now). Table III is taken from Shaw's paper and is her model for Dak$^{h}$ota Lexical Phonology. The Underlying Representation goes through three levels of morphology and phonology before exiting the Lexicon and entering the Post-lexical Phonology. It is in the Post-lexical Phonology that the derived word becomes phonetically realized. In Dak$^{h}$ota, the voicing of stops occurs in the Post-lexical Phonology. Processes of assimilation, dissimilation, lenition and fortition generally occur post-lexically.

Lexical Phonology deals with distinctive features or phonemes. Post-lexical Phonology deals with allophones. In English, the difference between aspirated and unaspirated voiceless stops is taken care of post-lexically since these are not distinctive features in English. Likewise, the devoicing of /l/ after a voiceless stop occurs postlexically as in the word /play/.

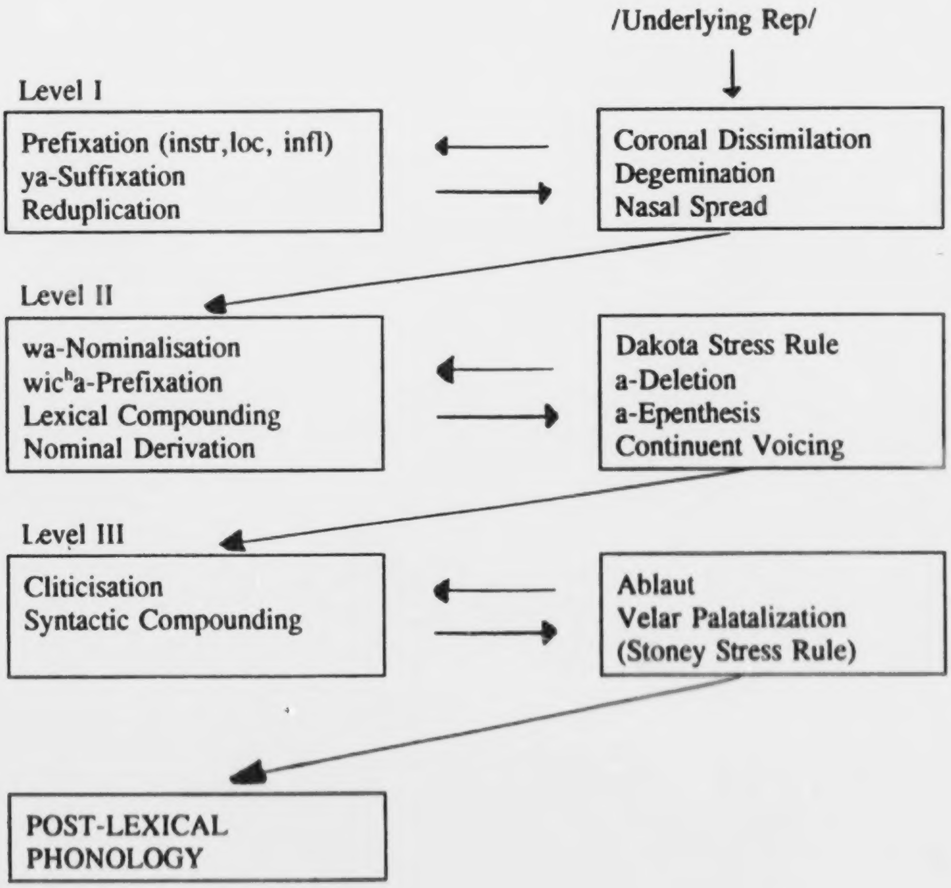

/Underlying Rep/
Level I
Prefixation (instr,loc, infl)
ya-Suffixation
Reduplication
Coronal Dissimilation
Degemination
Nasal Spread
Level II
wa-Nominalisation
wic$^{h}$a-Prefixation
Lexical Compounding
Nominal Derivation
Dakota Stress Rule
a-Deletion
a-Epenthesis
Continuent Voicing
Level III
Cliticisation
Syntactic Compounding
Ablaut
Velar Palatalization
(Stoney Stress Rule)
POST-LEXICAL PHONOLOGY

**TABLE III**: (from Shaw (1985:175) **Model for Dakota Lexical Phonology**

## DAK$^{h}$OTA MORPHOLOGICAL PROCESSES

The basic underlying element in Dak$^{h}$ota Morphology is the root. The root can go through several different word-formation (morphological) processes. Attached to the root can be prefixes, which includes instrumental and locative markers, nominalizers, and personal pronoun agreement affixes; and suffixes or enclitics, which can express temporal aspects, plurality, negation, gender of the speaker, or the type of speech act. Usually, the locative prefix(es) are ordered before the instrumental prefix(es). Pronominal affixes are usually next to the verb root but the position can vary depending on how closely the other prefixes are associated with the verbal meaning. Enclitics have a fairly rigid ordering which is described in Rood & Taylor(1976).

Nearly all prefixes take the (+) boundary which means they are added at Level I (the two noted examples are the nominalizer *wa-* and the third person plural animate pronoun *wicʰa-* which are added at Level II). Pronominal affixes come before the verbal root although their placement with respect to instrumental and locative affixes may vary. The first person singular and second person affixes are different for active and stative verbs. There is no third person marker except for the collective plural form. Plural forms are generally marked by the enclitic *=pi* although third person inanimate plural subjects are marked by the reduplication of the verb root.

(1) STATIVE

/wašte/ to be good,pretty

| | |
|---|---|
| mawášte | I am good |
| niwášte | you are good |
| wašté | he/she is good |
| uwášte | you and I are good |
| uwáštepi | we are good |
| niwáštepi | you(pl)are good |
| waštépi | they(anim,distr) are good |
| wicʰáwašte | they(anim,coll) are good |
| waštéšte | they(inanim) are good |

/oluluta/ to be sweltering (to feel hot and sweaty)

| | |
|---|---|
| omáluluta | I'm sweltering |
| oníluluta | you are sweltering |
| olúluta | he/she/it is sweltering |
| ukóluluta | you and I are sweltering |
| ukólulutapi | we are sweltering |
| onílulutapi | you(pl) are sweltering |
| olúlutapi | they(anim,dist) are sweltering |
| owícʰaluluta | they(anim,coll) are sweltering |

(2) ACTIVE

| /icu/ | to take |
|---|---|
| iwácu | I took (it) |
| iyácu | you took (it) |
| icú | he/she took (it) |
| ukícu | you and I took (it) |
| ukícupi | we took (it) |
| iyácupi | you(pl) took (it) |
| icúpi | they took (it) |

| /inayą/ | to have as a mother |
|---|---|
| ináwaye | I have her as a mother |
| ináyaye | you have her as a mother |
| ináye | he/she has her as a mother |
| iná?uye | you and I have her as a mother |
| iná?uyąpi | we have her as a mother |
| ináyayąpi | you(pl) have her as a mother |
| ináyąpi | they have her as a mother |

The pronominal affixes given above for the active verbs can be viewed as agentive affixes while the pronominal affixes given in the stative paradigm are patient affixes. In the two active verb conjugations given above the patient has been the third person singular which is unmarked in Dak$^h$óta. When the patient is not the third person singular the appropriate patient affix is used and precedes the agent prefix except in the case of the second person patient(singular and plural). The prefix *c*$^h$*i* is used for the forms which are equivalent to the English *I (verb) you*, and the first person plural agent precedes the second person patient form (singular and plural).

Additional Morphological rules include reduplication, and two types of compounding: lexical and syntactic. Reduplication of verbal roots serves several functions in Dak$^h$ota. It can mark the plurality of an inanimate subject, a repetitive action, intensification, and a distributive action. The actual process of reduplication consists of the copying of the final syllable of the root. It is important to note here that, underlyingly, there are two types of roots: consonant final (C#) and vowel final (V#). Although they each will surface as vowel final due to what Shaw (1986) calls the rule of a-Epenthesis (she calls it Stem Formation in Shaw (1980)). This rule adds a final vowel to the C# roots but only after reduplication has taken place. The forms of the possible underlying roots are shown here:

(3) V#

$(C_1^2V)C_1^2V$

| | |
|---|---|
| /niya/→ niyá | he breathes |
| /kte/→ kté | he kills (it) |
| /p$^h$a/→ p$^h$á | it barks |
| /naxma/→ naxmá | he hides (it) |
| /paha/→ pahá | hill |
| /yuγą/→ yuγą́ | he opens (it) |

(4) C#

$C_1^2VC$

| | |
|---|---|
| /caγ/→ cáγa | it freezes |
| /kaγ/→ káγa | he makes it |
| /top/→ tópa | it is four |
| /sap/→ sápa | it is black |
| /šap/→ šápa | it is dirty |
| /šųk/→ šų́ka | dog |

It should also be noted that the stress falls on the first syllable for C# roots and on the second syllable for V# roots of more than one syllable. This is accounted for by having a stress rule apply before the a-epenthesis rule. The Dakota Stress Rule (Shaw (1985)) places the stress on the second syllable of a word. If there is only one syllable, it is stressed.

(5) Dakota Stress Rule(DSR):

$$V \rightarrow \acute{V} \,/\, [(C_0V)C_0\_$$

Thus a monosyllabic word will be stressed and a word of two or more syllables will have stress on the second syllable. The DSR needs to apply after prefixation takes place since the prefixes can be stressed if they occupy the second syllable.

| (6) | | |
|---|---|---|
| | ksá | he cut it |
| | waksá | he cut it (*wa*- Absolutive) |
| | wakíksa | he cut his own |
| | wakíciksa | he cut it for him |
| | wayéciksa | you cut it for him |
| | wamíyeciksa | you cut it for me |
| | wawíc$^h$ayeciksa | you cut it for them |
| | wawáwic$^h$ayeciksa | you cut s.t. for them |

We should note that Prefixation comes before the DSR which comes before a-Epenthesis: Prefixation > DSR > a-Epenthesis. This fact is accounted for in

Shaw's (1985) Lexical Phonology model. She places Prefixation at Level I (as a morphological rule) and the DSR precedes a-Epenthesis at Level II (as phonological rules). She also places Reduplication at Level I since it appears to happen before the DSR. The relative ordering of Prefixation and Reduplication in Level I doesn't seem to matter since only the final syllable is reduplicated.

(7) **Reduplication of V# and C# roots**

V#

| | |
|---|---|
| /p$^h$e/ → p$^h$e+p$^h$é | 'are sharp' |
| /icu/ → icú+cu | 'pick up' |
| /wac$^h$i/ → wac$^h$í+c$^h$i | 'to dance' |
| /yamni/ → yamní+mni | 'three' |
| /háska/ → háska+ska | 'are tall' |

C#

| | |
|---|---|
| /xap/ → xap+xápa | 'to rustle' |
| /nųp/ → nųp+nų́pa | 'two' |
| /sap/ → sap+sápa | 'be black' |
| /kaγ/ → kax+káγa | 'to make' |
| /nak/ → nak+náka | 'to twitch' |

There are several things to take note of from the preceding examples. I have shown the underlying root (in slashed lines //) and the reduplicated form as it would appear after all Lexical processes (i.e. DSR, and a-Epenthesis for C# roots). The form /háska/ is written with the accent in the underlying form since the stress is always on the first /a/. Kiparsky(1982) considers a lexical entry to be a type of rule. His Elsewhere Condition states that a more specific rule will apply before a general rule and in effect block the general rule. Since /háska/ is already marked for stress, it will block the DSR.

Below is a derivation of a C# root (the syllable created by the reduplication process is shown in italics):

(8)

| | |
|---|---|
| Underlying form: | /sap/ 'to be black' |
| Level I | |
| Reduplication | [[sap][*sap*]] |
| Level II | |
| DSR | [sapsáp] |
| a-Epenthesis | [[sapsáp]a] |
| Surface form: | sapsápa |

The derivation of a V# root would appear as follows:

(9)

| | | |
|---|---|---|
| Underlying form: | /yamni/ | 'three' |
| Level I | | |
| Reduplication | [[yamni][*mni*]] | |
| Level II | | |
| DSR | [yamnímni] | |
| Surface form: | yamnímni | |

The next type of morphological process we need to look at is compounding. As we see by Shaw's model, there are two types of compounds: Lexical compounds at Level II and Syntactic compounds at Level III. As can be predicted, the differences between the two types of compounds will be seen in the stress patterns and also the presence or absence of the epenthetic *-a*. Lexical compounds which contain a C# root will lack the epenthetic *-a*. Also, since Lexical compounds are formed before the DSR, they will have only one stressed syllable while the Syntactic Compounds will have two. The Syntactic Compounds will have the epenthetic *-a*. An excellent example of the differences between Lexical and Syntactic Compounds are the words: $c^h$exzí ('brass kettle') and $c^h$éγazì ('yellow kettle'). The Lexical Compound ($c^h$exzí) has only one stressed syllable while the Syntactic Compound has two. The second stressed syllable has secondary stress. The example is composed of the morphemes {$c^h$ex} which means 'kettle' and {zi} which means 'yellow.' The following diagram shows their derivations:

(10)

| | Lexical | Syntactic |
|---|---|---|
| Underlying form: | /$c^h$ex/ /zi/ | /$c^h$ex/ /zi/ |
| Level I | --- | --- |
| Level II | | |
| Lex.Cmpnd | [[$c^h$ex][zi]] | --- |
| DSR | [[$c^h$ex][zí]] | [$c^h$éx] [zí] |
| a-Epen | | [[$c^h$éx]a] [zí] |
| Voicing | | [[$c^h$éγ]a] [zí] |
| Level III | | |
| Syn.Cmpnd | | [[$c^h$éγa][zí]] |
| Surface form: | $c^h$exzí | $c^h$éγazì |

## SYLLABLE STRUCTURE

A recent proposal by Kellogg (1991) attempts to simplify Shaw's (1986) phonological analysis by employing a prosodic theory of syllable structure. According to Kellogg, by using certain universal principles of syllable structure, one can do away with some of the rules from Shaw's analysis. I will first give a brief overview of the syllable theory, then I will show how Kellogg applies it to Dak$^{h}$ota and the rules it should replace, and finally I will show how the method falls short of achieving its simplification.

According to Kellogg, within the lexicon, the Lak$^{h}$ota syllable is open. The concept of Structure Preservation is upheld within the lexicon. Structure Preservation does not apply Post-lexically meaning that syllable codas can (and do) exist there. Any Lak$^{h}$ota syllable in the lexicon will adhere to the syllabic template: C$_0$V. A word final consonant will be regarded as extraprosodic (ex). Extraprosodic material can only exist word-finally. If an extraprosodic unit which ends up in word-internal position due to some morphological process (such as reduplication) it must either associate with the onset of the following syllable or be erased by the process of Stray Erasure. There are two things that can happen to the word-final extraprosodic material: i) will either form the onset of a new syllable or ii) if it occurs at the last level of the lexicon it may become a coda post-lexically.

(11) **Rules for Creating Syllable Structure**(Kellogg, 1991:32):

a) Moraify all sonorous segments that are [-cons].

$$\text{a) } V \rightarrow \begin{array}{c}\mu\\ |\\ V\end{array}$$

b) Project a syllable node over each mora.

$$\text{b) } \mu \rightarrow \begin{array}{c}\sigma\\ |\\ \mu\end{array}$$

c) Associate all licensable onsets to syllable nodes.

$$\text{c) } C \rightarrow \begin{array}{c}\sigma\\ |\\ C\end{array}$$

d) Assign extraprosodicity to all word-final consonants.

$$\text{d) } C]_w \rightarrow \overset{\text{[ex]}}{\overset{|}{C}}]_w$$

Below are some of the phonological rules given by Shaw (1980, 1985) which Kellogg claims we can replace by Prosodic Theory:
(12)

a. **Degemination**:

$$C_i \rightarrow \emptyset \, / \, \_\_ + C_i$$

/k$^h$ak/ k$^h$ak-k$^h$ak →[k$^h$ak$^h$áka]
'to rattle'
/xux/ xux-xux →[xuxúγa]
'to thunder'

b. **Cluster Simplification**:

$$C \rightarrow \emptyset \, / \, \_\_CC$$

/xpec/ xpec-xpec →[xpexpéca]
'lifeless'
/ksap/ ksap-ksap →[ksaksápa]
'be wise'
(lex-compd) [p$^h$et-snis](fire-to fade)→p$^h$esníza
'embers'

c. **Dissimilation**:

$$\begin{bmatrix}\text{-cont} \\ \text{+cor}\end{bmatrix} \rightarrow \begin{bmatrix}\text{-cor} \\ \text{-ant} \\ \text{-son}\end{bmatrix} / \, \_\_ + [\text{+cor}]$$

/šic/ šic-šic →[šikšíca]
'be bad'

d. **Epenthesis**:

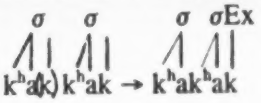
$$\emptyset \rightarrow a \, / \, C\_\_ \left\{\begin{matrix} = \\ \# \end{matrix}\right\}$$

/cap/ cap-a →[cápa] 'beaver'

Degemination[1] and Cluster Simplification can be accounted for by the

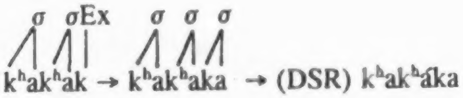
Stray Erasure of the outermost potential onset which cannot associate with the following syllable due to well-formedness conditions of the onsets. The set of possible onsets in Dakʰota is given in Table IV below. In the example above for /kʰak/, the final *k* cannot associate to the onset of the next syllable for the cluster *kkʰ* is not permissible so it is deleted.

(13)



The final extraprosodic *k* in *kʰakʰak* becomes the onset for a new syllable with the addition of an *-a*. This accounts for the rule of a-Epenthesis.

(14)



Only word-final consonants can be extra-prosodically licensed, thus non-word-final consonants are deleted by Stray Erasure. 'Given the proposed open syllable structure, all consonants must be syllabified as onsets, with the exception of those occurring word-finally (Kellogg (1991:35)).'

| | p | t | k | s | š | c | l | n | m | w |
|---|---|---|---|---|---|---|---|---|---|---|
| p | | pt | | ps | pš | pc | | | | |
| t | | | tk | | | | | | | |
| k | kp | kt | | ks | kš | kc | kl | kn | km | kw |
| s | sp | st | sk | | | sc | sl | sn | sm | sw |
| š | šp | št | šk | | | šc? | šl | šn | šm | šw |
| x | xp | xt | | | | xc | xl | xn | xm | xw |

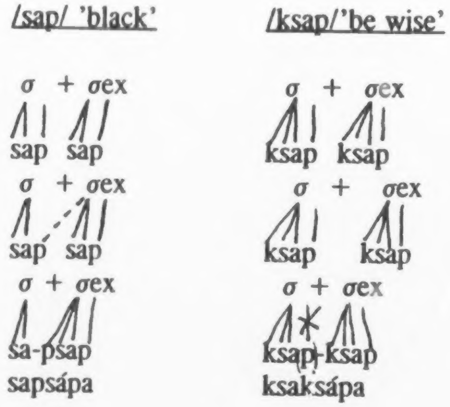
**TABLE IV**: (from Shaw (1989:7))Possible Syllable Onsets

Using Kellogg's syllable analysis, the reduplicative template consists of the final syllable of the root plus any extra prosodic consonant (a C# root has an extraprosodic consonant) copied to the right of the root. For V# roots this means that the final syllable is repeated. For C# roots the final syllable plus the extraprosodic final consonant are copied. The original final consonant can no longer be viewed as extraprosodic (since only 'word' final consonants can be licensed as such) and must either associate to the onset of the duplicate syllable or be deleted by the process of Stray Erasure.

(15)

Process

a. affixation of the reduplicating template:
b. copy melody:
c. association:
(+Onset Rule)
d. Stray Erasure:
e. final form:



The processes of association and Stray Erasure are viewed as universal conventions. They do not need to be listed as rules for a language but they do have to follow the constraints of a language's syllable structure.

The theory seems to work well for reduplicated forms but not for Lexical Compounds. Kellogg (1991:38) claims that Stray Erasure applies at the end of each level of the lexicon. This would mean that there should be no geminates at any level. This contrasts with Shaw's analysis. Shaw differentiates between Levels I and II by showing that geminates don't occur at Level I but can occur at Level II. She gives the examples of the Lexical Compounds (Shaw 1985:185):

| (16) | [c$^{h}$ap][p$^{h}$at] | c$^{h}$app$^{h}$áta | 'butcher beavers'(beaver+butcher) |
|---|---|---|---|
| | [wat][t$^{h}$ete] | wadt$^{h}$éte | 'gunwale' |
| | [t$^{h}$ok][k'u] | t$^{h}$okk'ú | 'to give over an enemy' |

Boas and Deloria (1941:13) also list several compounds where there appears to be gemination[2]

| (17) | hąppáhi | 'to collect moccasins'(moccasin+collect) |
|---|---|---|
| | hąpp$^{h}$áxta | 'to tie moccasins in a bundle'(mocc.+tie in bundle) |
| | šųkkáštaka | 'to whip a horse'(horse+to whip) |

| | |
|---|---|
| íyotakk$^h$iya | 'to make sit down' |
| šųkkoyakya | 'to rope a horse' (from Buechel(1970)) |

Either the above examples are exceptions or Kellogg's analysis doesn't hold up. I hesitate to call these exceptions simply because they are few in number. The conditions needed to form possible geminates limit the number that can be formed. A C# root (usually a noun) must form the first member of a compound and the second member must begin with the same consonant. Add to this the fact that only a limited number of phonemes actually occur root finally; and we narrow the possibilities more. It would be better for a theory to be able to explain these clusters rather than to list them as exceptions. But geminates are not the only consonant clusters that appear. There are other Lexical Compounds which contain complex consonant clusters which are not acceptable syllable onsets.

| | |
|---|---|
| šųkská | 'white horse'(horse+white) |
| p$^h$elmná | 'smell of fire'[p$^h$et][mna] (fire+smell) |
| capkté | 'to kill beavers' (beaver+to kill) |

It becomes clear that there is a difference in how the formation of complex is handled in Reduplication and Lexical Compounding. Reduplicated forms seem to adhere to strict well-formedness rules and Structure Preservation which quickly delete or change any segments which could make an unacceptable onset. Lexical Compounding rules seem to be more lax. We could possibly temper Kellogg's analysis by limiting Stray Erasure to Level I. This would handle the dilemma but then it would be hard to argue that her analysis is any simpler than Shaw's. And what about the condition given by the first Lexical Phonologists (Kiparsky) that the domain of Structure Preservation is the Lexicon? More recent work done in Lexical Phonology has shown that many of the strict conditions put on the theory in its formative years no longer hold up. In the introduction to Studies in Lexical Phonology, Kaisse and Hargus (1993:16) write that 'with nearly a decade of subsequent work, we now know that many of these characteristics **(such as Structure Preservation)** cannot be considered diagnostic of the lexical or postlexical status of a rule.'(Bold letters are my addition). Also: 'In some languages, structure preservation appears to hold of postlexical rules, whereas in other languages, some rules which are clearly lexical (albeit word-level) may not be structure-preserving (Kaisse and Hargus (1993:16)).'

The major drawback in limiting Structure Preservation to Level I is that the structural rule of epenthesizing a root final *-a* to C# roots occurs at Level II after Lexical Compounding (Lexical Compounds formed from C# roots do not have the epenthesized -a but Syntactic Compounds (Level III) do). I do not claim to have the answer to this problem but it reveals that more work is needed in the area of Dak$^h$ota syllable structure.

## NOTES

1.I believe I need to make some comments about geminates in Dak$^h$ota. I will be the first to admit that on the surface (or phonetically) there don't appear to be geminates (i.e. two identical segments). But underlyingly they can arise! As in the example given /xux/ will reduplicate to /xux+xux/. I consider the two adjacent x's to be geminates although they are quickly destroyed by the Stray Erasure of the left-most one since it cannot associate to the following onset giving /xu+xux/. I will call them geminates for now although perhaps we should call them *virtual* geminates.

2.Boas & Deloria (1941:13) do not write the doubled consonants as I have done. They mark the consonant in question with a [˙] to show the extra length of the segment. For example, where Boas & Deloria write *hąp˙áhi*, I write *hąppáhi*.

## REFERENCES

Buechel, Rev. Eugene (1970). *Lakota-English Dictionary*. Red Cloud Indian School, Pine Ridge, South Dakota.

Boas, Franz & Ella Deloria (1941). *Dakota grammar*. Memoirs of the National Academy of Sciences 23. Part 2.

Kaisse, Ellen M. and Sharon Hargus (eds) (1993). Introduction. *Phonetics and Phonology: Studies in Lexical Phonology vol 4*. Academic Press, inc. San Diego.

Kaisse, Ellen M. & Patricia A. Shaw (1985). On the theory of Lexical Phonology. *Phonology Yearbook*. 2. 1-30.

Kellogg, Kim. (1991). The Role of Extraprosodicity in Prosodic Reduplication:

A case study of reduplication in Lakota. *Linguistic Notes from La Jolla.* v16 29-55

Kiparsky, Paul. (1982). Lexical morphology and phonology. In I.s. Yang (ed.) *Linguistics in the morning calm.* Seoul: Hanshin. 3-91.

Kiparsky, Paul. (1985). Some consequences of Lexical Phonology. *Phonology Yearbook.* 2. 85-138.

Mohanan, K.P. (1986). *The theory of Lexical Phonology.* D.Reidel Publishing, Dordrecht, Holland.

Rood, David & Allan Taylor (1976). *Beginning Lakhota.* University of Colorado Lakhota Project: University of Colorado Department of Linguistics.

Shaw, Patricia A. (1980). Theoretical Issues in Dakota Phonology and Morphology.

Shaw, Patricia A. (1985). Modularism and substantive constraints in Dakota Lexical Phonology. *Phonology Yearbook*.2. 173-202.

Shaw, Patricia A. (1989). The Complex Status of Comples Segments in Dakota. *Theoretical perspectives on Native American Languages.* SUNY Press, Albany.

# NOTES ON THE KANSA WORD LIST OF MAXIMILIAN, PRINCE OF WIED[1]

Robert L. Rankin
University of Kansas

> 'It is the duty of every traveller in distant, little-known lands to make contributions, according to his ability, to the knowledge of their languages....'
> Maximilian zu Wied-Neuwied

Abstract: The Kansa word listed collected in ca. 1833 by Prince Maximilian of Wied is compared with two more recent transcriptions of the same terms. Observations about the phonology are offered which, in turn, clarify the ultimate source of Maximilian's list.

One of the earliest Kansa language vocabularies, if not *the* earliest, in existence is that collected by Alexander Philip Maximilian, Prince of Wied-Neuwied, during his voyage through the plains of North America between 1832 and 1834. The list was published as an appendix to Maximilian's journals in 1843, along with vocabularies of a large number of other plains languages. In this paper the words on Maximilian's 1833 list are compared with the same Kansa words recorded by the rev. James Owen Dorsey, probably sometime between 1888-1890 and by this author in the 1970's.[2] From this comparison it is possible to make a few linguistic observations that seem pertinent, primarily about the phonology of the words on Maximilian's list, then, based on these and other observations, certain inferences can be drawn about the source of the list itself.

The Maximilian and Dorsey lists are presented below in their respective authors' original notation; the Rankin list is surface phonemic (with the exact status of vowel length still undetermined). Comments offered by the several authors are placed beneath the pertinent entries in the proper columns.

| | Maximilian | Dorsey | Rankin |
|---|---|---|---|
| Arrow | mah | ma | ma |
| Bear (black) | uassóbä | wa-sa´-be | wasábe |
| Child | schɨnga-schɨnga | jiñ-ga´ jiñ´-ga | hɨgá hɨga |

*Kansas Working Papers in Linguistics,* Volume 19, Number 2, 1994, pp. 201-210

| | | | |
|---|---|---|---|
| Earth | móhnika | ma$^{n}$-yiñ´-ka | mą́(y)ɨ̨kka |
| | (*n* nearly like *h*) | | |
| Eye | ischtá | i-cta´ | ištá |
| Fire | pähdjé | pye´-dje | ppé·ǰe |
| | (*j* French) | | |
| God | wahkondagä | wa-kan´-da-gi | wakką́dagi |
| | | (*conjurer*) | (*doctor*) |
| Hair | pa-hɨ̨ | pa-hü´ | ppahü´ |
| Hand | nom-pő | na$^{n}$-be´ | nǫbé nąbe |
| | (*om* French) | nu$^{n}$-be´ | |
| Head | pah | pa (nose) | ppa |
| Island | rumätschɨ | ni-ba´-se | nibáse |
| Man | niká | ni´-ka | níkka |
| Mountain | pahű | ba-xu´ | (bahé) |
| Mouth | hüh | i´-ha | íha |
| Pipe (tobacco) | nah-hi-ba | na´-nü-u$^{n}$´-ba | nąnǫ́ba |
| River | watɨschka | wa-tci$^{n}$´-cka | wačhíška |
| | (*t* often like *h*) | (archaic) | |
| Sun | pih | mi$^{n}$ | mi, mɨ̨ |
| Tomahawk | ma-sospä-jingá | ma$^{n}$´-hi$^{n}$-spe | mą́·hɨ̨spe |
| | (*j* French) | +jiñ-ga | +hɨ̨ga |
| Water | nih | ni | ni, nɨ̨ |
| Woman (wife) | wah-ko | wa-k'u´ | wakʔó |

Accompanying his word list Maximilian wrote briefly of the Kansa people and the affinities of the Kansa language:

> The Konsas, or Kansas Indians have always lived on the river of that name (Gallatin, p. 127). For the past thirty years they have lived at peace with the Osages; and the two tribes have intermarried. They still number fifteen hundred souls, and possess a tract of three thousand acres. They speak the Osage dialect, which belongs to the Dacóta linguistic group. (Thwaites, vol. 24, p. 229)

And although we find such delightful comments as 'Most of the... languages of the Missouri Valley can more easily be reproduced by the Germans and Dutch than by other nations; because,... their own speech abounds in gutturals and regularly has hard endings,' in point of fact, Maximilian took special care to write the words he heard accurately. He had read Duponceau and Pickering carefully and, citing both, discusses problems of orthography at some length.

He used what amounts to a transcription system based on a combination of German and French orthographic norms with the addition of a number of special diacritics and liberal use of explanatory notes. His system utilizes primarily German spelling conventions, resorting to French only when German lacks a particular sound. Thus, for example, nasal vowels and [ž] are written the French way, i.e., with postposed nasal consonants and the letter *j*, respectively.

Turning to Maximilian's Kansa list, a number of things can be said.

The precise status of vowel length in Kansa is still a problem today. Recordings made in the 1970's show it sporadically in words where it occurs regularly in other Siouan languages such as Winnebago or Crow. So, while undoubtedly significant at one time, it apparently became variable as use of the language decreased and presented difficult problems for the transcriber in later years. Maximilian took special note of vowel length in all of his word lists stating, 'I have tried to indicate the length of a syllable by adding an *h*,...' (Thwaites, vol. 24, p. 206) Maximilian adopted the device used in German orthography for noting long vowels, i.e., his use of *h* does not parallel attempts English speakers sometimes make to indicate continental vowel values. We find length indicated in the following items: ARROW, HEAD, MOUTH, SUN, WATER, EARTH, FIRE, GOD, PIPE, WOMAN. This list suggests two generalizations.

First, Maximilian notes the vowels in all monosyllables as long. Monosyllables are indeed all long in a number of Siouan languages. White Eagle and Miner (personal communication) both report this independently for Winnebago, and it also seems to be true of at least Chiwere and Crow in addition. Dorsey's (ca. 1888) normalized Kansa transcriptions do not give clear indication of vowel length, so they are of little help to us, and length is subject to considerable variability in today's rapidly obsolescing Dhegiha languages (Rankin, field notes; Richard T. Carter, John E. Koontz, personal communication). So we have not known how closely Dhegiha followed common Siouan lengthening rules, and it would be nice if Maximilian's transcriptions provided the evidence. That Maximilian wrote length for all of the Kansa monosyllables on his list (ARROW, HEAD, MOUTH, SUN, WATER), appears to confirm that the rule did apply generally in Kansa in the 1830's, but we shall have to return to this problem presently.

Second, in Chiwere, and with considerable variation in Dhegiha today, all accented initial syllables appear to be long. For the most part Maximilian's list shows this also (EARTH, FIRE, PIPE). The only exception is CHILD, which should have length but where none is indicated. GOD and WOMAN perhaps should not have the long vowels attributed to them in the word list.

Voicing of obstruents is another phenomenon that is important as it is one of the phonological features that differentiates Kansa from Osage. In the Kansa of the 1970's all instances of the lax (i.e., unaspirated, unglottalized and ungeminated) series of stops had voiced (except, of course, in clusters with voiceless fricatives), i.e., common Dhegiha **p t k* > *b d (ǰ) g*.

In the century-old Dorsey materials there are exceptions to this voicing. Dorsey wrote numerous instances of [k] for modern [g]. There are fewer instances of [t] for modern [d], and I was able to find no instances at all of [p] for modern [b]. This parallels the situation in other Dhegiha languages that I have discussed in greater detail elsewhere. Within Dhegiha the lax series has voiced most completely in Omaha-Ponca and Kansa, partially in Quapaw and scarcely at all in Osage. The philological and comparative evidence indicates that the voicing began with the bilabials and proceeded through the dentals affecting the velars last.

The Maximilian list shows no voiced stops where any later source has voiceless stops, but it does have a few apparently conservative cases of voiceless stops where the later sources have innovated voicing in the words for HAND and MOUNTAIN. Interestingly, these involve labials, all of which had voiced in the later sources available to us. All are pretonic; in the Maximilian list all posttonic lax stops are written voiced. Among the fricatives there is one substitution of *š* for *ž* in CHILD, which is written with *sch*; in TOMAHAWK the same morpheme, *žįga* SMALL (unstressed) is written with the letter *j* representing *ž*.[3]

Beyond the voicing discussed above, a few additional observations regarding Maximilian's transcription of obstruent consonants are possible.

There is no graphic sign of the glottalization in WOMAN, but the word, written as it is with *k*, shows at least that common Dhegiha **xʔ* (preserved still in Quapaw) had already merged with *kʔ* by the 1830's, i.e., **waxʔó* > Kansa *wakʔó*. Lack of a graphic representation of glottalization is not surprising however, as even the most talented amateur phoneticians who attempted to record American languages often had trouble deciding what to do with it.

Only one aspirate is present in the data. RIVER carries Maximilian's cryptic comment '*t* often like *h*', but while he could have been referring to the aspiration of the *t*, which is phonemically *th* throughout the Dhegiha Siouan subgroup, it seems just as likely that he was referring to affrication instead (Kansa *th* > [čh] regularly preceding front vowels). Interestingly, the sound is nonetheless written by Maximilian as *t*, not the perhaps to be expected *tsch*. That affrication of dentals was proceeding apace however, is shown by FIRE where **ppé·te* has already become *ppé·ǰe*.

Tense or geminate stops are always written by Maximilian with letters representing voiceless stops, but otherwise they are not distinguished from the other stop series. In other words, Maximilian did not distinguish aspirates from tense (geminate) or glottalized stops, hardly surprising since only the one aspirate and one ejective were represented in his Kansa list.

Maximilian (Thwaites, vol. 24, p. 206) states, 'As the French acute accent seemed to me entirely adequate for indicating the correct intonation of the Indian words, I have chosen it to show on what letter or syllable the emphasis is to be placed. Occasionally, owing to haste or lack of time, it may have been forgotten or omitted.' The translator mentions (*ibid.* p. 209) that 'The vertical accent seems to be used to indicate secondary, or less emphatic stress....'

Little can be said about accent in this word list other than that it generally seems to correspond to the notation of stress in the other sources. FIRE and MAN appear mistranscribed, or misrepresented by Maximilian's source, as stress falls on the initial syllable of these words in Kansa and throughout Dhegiha.

The spelling of EARTH suggests that the development of modern Kansa *y* from common Dhegiha **r* in a nasal environment may not have been complete and that some occlusion may still have been present. In this word Ponca and Quapaw have *n* while Osage has *δ* and Kansa has *y*. The earlier stage in Kansa probably corresponded closely to the Osage. Maximilian's comment '*n* nearly like *h*' may refer to lack of occlusion, i.e., a [δ]-like phone preceding a nasal vowel.

The final vowel of HAIR is written *i* despite the fact that there was a readily available symbol for *ü* (cf. MOUNTAIN). Common Dhegiha *ü* is preserved yet today in both Kansa and Osage, but for at least one hundred years there has been a certain amount of phonetic fluctuation when it comes to rounding. HAIR, if rendered correctly to, and recorded by, Maximilian, could be interpreted as evidence for similar fluctuation as early as the 1830's.

Finally, there are several words that are either unrecognizable or appear quite different from their later forms.

MOUTH should not have a rounded vowel and cannot be explained except perhaps as a copying error. It does not have a rounded vowel in any Siouan language, so this is one instance in which access to the original manuscript would clearly be very useful. There must be an editor's error or misprint here.

PIPE also almost certainly contains a copying or printing error. It was probably recorded *nah-ni-ba* by Maximilian with later misinterpretation of the second *n* as *h* (Kansa 'tobacco' is *nąní* varying with *nąnũ*).

RIVER shows oral *í* rather than the aberrant *į́* recorded by Dorsey. This not only agrees with what this author found in the 1970's but with the oral vowel found in this word in all the rest of the Dhegiha languages also. Here it was probably Dorsey who erred.

We may now consider briefly the source of Maximilian's Kansa vocabulary. In most instances he indicated the source of each vocabulary in a footnote; unfortunately Kansa is one of the very few cases in which no source is given, at least in Thwaites. About his sources generally, Maximilian writes:

> I have written these vocabularies, in part from the pronunciation of the Indians themselves; in

> part, from that of interpreters, who are usually half-breeds, and therefore thoroughly acquainted with the Indian as well as with the French or English languages--they have, at least, lived for a long time with those nations; finally, in part from the pronunciation and with the kind assistance of the Indian agent, Major Dougherty, who speaks several of these languages fluently. (Thwaites, vol. 24, p. 207.)

Major John Dougherty was Indian agent at the Bellevue (Nebraska) agency just south of Omaha. He accompanied Maximilian up river to Bellevue by steamer from near St. Louis. He was primarily in charge of relations with the Omahas, Otoes and Pawnees (all Nebraska tribes) according to Maximilian. In one incident described he also has dealings with a group of Ioways. The Kansas, on the other hand, had a sub-agent with whom they dealt at Leavenworth Cantonment (Fort Leavenworth after 1832).

William E. Unrau, in his 1971 history of the Kansa tribe, mentions that Dougherty visited the Kansas on only two documented occasions, and was, in fact, accused by a prominent official of the Missouri Fur Trading Co., Joshua Pilcher, of having neglected them. The Kansa at this time were living about ninety miles west of Leavenworth, which was situated on the Missouri River just north of Westport (today a Kansas City neighborhood). Nevertheless, despite his supposedly brief contact with them, Dougherty appears to be the probable source of Maximilian's Kansa word list.

Whoever gave him the list seems to have reproduced Kansa fairly well for the most part, but in a few cases he apparently could not recall the correct form, and when this happened he always gave the cognate or other analog as it occurs in one of the Nebraska Siouan languages, Otoe or Omaha-Ponca. This happens at least three times, in the words for ISLAND, SUN and TOMAHAWK.

Below I have reproduced Maximilian's Kansa list again, this time compared with his Otoe word list, a list specifically obtained from Major Dougherty.

| English | Kansa | Otoe |
|---|---|---|
| Arrow | mah | míto |
| Bear (black) | uassóbä | montchâ |

| | | |
|---|---|---|
| Child | schínga-schínga | tchitching-ä |
| Earth | móhnika | |
| Eye | ischtá | ichtá |
| Fire | pähdjé | pedjé |
| God | wahkondagä | wahkonda |
| Hair | pa-hí | pa-hí |
| Hand | nom-pó | nau-uä |
| Head | pah | pa |
| **Island** | **rumätschí** | **rú-mi-tschí** |
| | | or **rumaetschi** |
| Man | niká | uong-gäh |
| Mountain | pahů́ | hämokschǻ |
| Mouth | hüh | hí |
| Pipe (tobacco) | nah-hi-ba | ra-no-wǻ |
| River | watíschka | nisch-nong-ä |
| **Sun** | **pih** | **píh** |
| **Tomahawk** | **ma-sospä-jingá** | i-sua-ing-ä |
| Water | nih | níh |
| Woman (wife) | wah-ko | ina-hakä |

On this list I have boldfaced the three suspect items.

ISLAND, Kansa *rumätschí*, Otoe *rú-mi-tschí* or *rumaetschi* cannot be completely reconciled with any later occurring Kansa form. In addition, [r] is not a Kansa sound, not even allophonically. The recorded form is virtually identical to Maximilian's Otoe entry however, which he says was 'written from the pronunciation of Major Dougherty,...who understands the language thoroughly.' (Thwaites, vol. 24, p. 285)

SUN should have initial *m*. This nasal is present in every Dhegiha language, but in Otoe it has the modern form *bi*, where *b* is phonetically either an unaspirated [p] or a [b]. Maximilian's Otoe entry for SUN is written *píh*, exactly like the suspect Kansa entry.

TOMAHAWK in Kansa and Osage is based on the word that is usually translated 'knife', *mą́hį* (thence *mą́hįspe*). Here, instead of containing *mą́hį*, it is compounded from *mąze-* 'metal' and the second element, *-spe*, from the later attested form, the word for 'axe'. This follows neither the Kansa nor Otoe pattern, but it is close to the modern Omaha *mązǫspe* 'axe'.[4] Recall that Dougherty's assignment involved both the Otoes and Omahas centrally.

The only other possible source of Maximilian's

Kansa list would have been Osage. He remarks more than once on the similarity of Kansa and Osage. For example on Kansa he writes (Thwaites, vol. 22, p. 252): 'Their language is entirely the same as that of the Osages, and the language of these two people is only a dialect, originally not different from that of the Omahas and Puncas, being distinguished only by the pronunciation, and not by its roots.'

Maximilian's Osage sample was 'written from the pronunciation of Mr. Chardon, who had lived a long time among the Osages and understood the language perfectly.' (Thwaites, vol 24, p.296) Osage was spoken far to the South in southwest Missouri and adjacent areas of Oklahoma and southeast Kansas. The fact that the non-Kansa substitutions on Maximilian's Kansa list are taken from Otoe and Omaha (of Nebraska) rather than the more closely related but geographically distant Osage all but rules out Chardon as a source for the Kansa list, leaving Dougherty the only likely candidate.

In spite of the fact that Maximilian's 1833 Kansa word list is clearly second-hand, most of the vocabulary is obviously Kansa and even contains examples of the features that separate Kansa from Osage. Observations on the progression of consonant voicing, affrication of dentals, nasalization and accent can be at least tentatively advanced.

Finally, it can be stated with some certainty that Major John Dougherty was the source of the Kansa vocabulary and that Dougherty's primary linguistic competence in Siouan was in Chiwere (Ioway-Otoe-Missouria) rather than Dhegiha (Kansa, Osage, Omaha-Ponca, Quapaw). The vowel length shown for all monosyllabic Kansa words then may be misleading however, since in English, monosyllables cannot end in short vowels and Dougherty would probably have pronounced them long no matter what he actually heard. Vowel length aside though, judging from the Kansa and Otoe lists, Dougherty appears to have had a fairly good ear for Siouan phonologies.

## NOTES

[1] This preliminary note is based on my analysis of the published versions of Maximilian's Kansa, Otoe and Osage word lists (Thwaites 1906, vol. 22-24) and so may include inaccuracies introduced by the editor

and/or publisher. I would have preferred to work from primary sources of course, but these are not yet available to me.

Over the long run it might also be quite useful to consult the papers of Maj. John Dougherty, who was the source of several of Maximilian's lists. Maj. Dougherty evidently had a good ear for the phonetics of the Siouan languages and may have recorded some of them more extensively himself. According to the Fort Leavenworth Archivist, Steve Allie (personal communication), Dougherty's papers may be stored at Carlisle Barracks in Pennsylvania.

[2] I am grateful to the American Philosophical Society and to the Graduate Research Fund of the University of Kansas, both of which supported my field work on the Kansa language. I am also especially grateful to Mrs. Maude Rowe and her family, without whose help the Kansa project could never have been undertaken. Mrs. Rowe worked patiently with me for several years beginning in 1974, and most of my transcription is based on her pronunciation of Kansa.

[3] In the modern language the unstressed allomorph serves mostly as a diminutive marker. In this form it has lost both its friction and oral stop, appearing in the variant *hįga* [hįŋə].

[4] The Omaha term itself is evidently a loanword from Santee Dakota, where it has the same shape. So, in fact, either language could have been the source.

## REFERENCES

Dorsey, James Owen. ca. 1888. Kansa slip file. National Anthropological Archives, Smithsonian Institution.

Rankin, Robert L. 1987. *Kansa Dictionary*. MS.

Thwaites, Reuben Gold, ed. 1906. *Travels in the Interior of North America by Maximilian, Prince of Wied*, in the series *Early Western Travels: 1748-1846*, vols. 22-24. Translated from German by Hannibal Evans Lloyd. Cleveland: Arthur H. Clark.

Unrau, William E. 1971. *The Kansa Indians*. Norman: University of Oklahoma Press.

Reflexives in Mohawk
Nancy Bonvillain

Word-Level Nominalization in Choctaw
Marcia Haag

Notes on the Kansa Word List of Maximilian, Prince of Wied
Robert L. Rankin

The Limit of Structure Preservation in Dakʰota Lexical Phonology
John Kyle

Proto-Algic V: Doublets and their Implications
Paul Proulx

Karankawa Linguistic Materials
Anthony P. Grant